

FIGURE 1

CGGCAGCCCTGACGTGATGAGCTCAACCAGCAGAGACATTCCATCCCAAGAGAGGTCTGCGTGACGCGTCCGGGAGG
 CCACCCCTCAGCAAGACCACCGTACAGTTGGTGGAAAGGGGTGACAGCTGCATTCTCCTGTGCCTACCACGTAACCAAA
AATGAAGGGAGAACTACTGTTACAAGCCGCCCTGGTGTGCCTGGCATGCTGTGCCACAGCCATGCCTTGCCCCAG
 AGCGGCGGGGGCACCTGCGGCCCTCCTCCATGGCACCATGAGAAGGGCAAGGAGGGCAGGTGCTACAGCGCTCC
 AAGCGTGGCTGGGTCTGGAACCAGTTCTCGTGTAGAGGAGTACACCGGGCTGACCCCGTGTGGGCAGGCT
 TCATTCAAGATATTGACTCTGGTGTAGGGAACATTAACATTCATTCTCAGGGAAAGGAGGCTGGAACCATTGGTGA
 TTGATGACAAATCAGGGAACATTCAATGCCACCAAGACGTTGGATCGAGAAGAGAGAGCCCAGTACACGTTGATGGCT
 CAGGCAGGTGGACAGGGACACCAATCGGCCACTGGAGCCACCGTCGAATTCAAGGTCCAGGACATTAATGA
 CAACCCCTCCGGAGTTCTGCACGAGACCTATCATGCCAACGTGCCTGAGAGGTCCAATGTGGAACGTCAGTAATCC
 AGGTGACAGCTTCAGATGCAGATGACCCACTTATGAAATAGCGCCAAGTTAGTGTACAGTATCCTCGAAGGACAA
 CCCTATTTCGGTGGAAAGCACAGACAGGTATCATCAGAACAGCCCTACCCAACATGGACAGGGAGGCCAAGGAGGA
 GTACCACGTGGTGTCCAGGCCAAGGACATGGTGGACATGGCGACTCTCAGGGACAACCAAAGTGCAGATCA
 CACTGACCGATGTCATGACAACCCACCAAAGTTCCGAGAGGCTATACCAGATGTCAGAACAGCCGTC
 CCTGGGGAGGAAGTAGGAAGAGTGAAGCTAAAGATCCAGACATTGGAGAAATGGCTTAGTCACATAAATATTGT
 TGATGGAGATGGTATGGAATCGTTGAAATCAAACGGACTATGAAACACAGGGGGGTGATAAAGCTGAAAAAGC
 CTGTAGATTTGAAACCGAAAGAGCCTATAGCTGAAGGTAGAGGCAGCCAACGTGCACATGACCCGAAGTTATC
 AGCAATGGCCCTTCAAGGACACTGTGACCGTCAAGATCTCAGTAGAAGATGTCAGTGTGAGCCCTATGTTCTGGC
 CCCAAGTTACATCCACGAAGTCCAAGAAAATGCAGCTGCTGGCACCGTGGTTGGAGAGTGCATGCCAAAGACCCCTG
 ATGCTGCCAACAGCCGATAAGGTATTCCATCGATCGTCACACTGACCTCGACAGATTTCACTATTAATCCAGAG
 GATGGTTTATTAAAACACAAACCTCTGGATAGAGAGGAAACAGCCTGGCTCAACATCAGTGTCTTGCAGCAGA
 AATCCACAATCGGCATCAGGAAGCCAAGTCCCAGTGGCATTAGGGCCTTGATGTCACAGATAATGCTCCAAGT
 TTGCTGCCCTTATGAAGGTTCATCTGTGAGAGTGTAGCAGACCAAGCCACTTCCAACCAGCAATTGTTACAATT
 AGTGCAGATGACAAGGATGACACGGCCAATGGACCAAGATTTATCTCAGCCTACCCCTGAAATCATTACAATCC
 AAATTTCACAGTCAGAGACAACCGAGATAACACACAGCAGGGCTGTACGCCGGCGTGGAGGGTCAGTCGGCAGAAC
 AGGACTTGTACCTCTGCCCATAGTGTAGCAGCGATGGCGGATCCGCCATGAGTAGCACCAACACCCCTACCCATC
 AAAGTCTCGGGTGCACGTGAACGGGCACCGCTCTGCACAGCAGAGGCCACATTCTGAACGCCGGCTGAG
 CACAGGCCCTGATGCCATCCTGCCATCGTCATTCTCCTGGCATTGAGTAGTATTGTTGTGACCCCTGAGAA
 GGCAAAAGAAAGAACCACTCATTGTCTTGAGGAAGAAGATGTCGTGAGAACATCATTACTATGATGATGAAGGG
 GGTGGGGAGAACACAGAACGCCCTTGATATTGCCACCCCTCCAGAACCTGATGGTATCAATGGATTATCCCCCG
 CAAAGACATCAAACCTGAGTATCAGTACATGCCAGACCTGGCTCCGGCCAGGCCAACAGCGTGGATGTCGATG
 ACTTCATCAACACGAGAACAGGAGGGCAGACAATGACCCACGGCTCCTTATGACTCATTCAAATCTACGGT
 TATGAAGGCAGGGCTCAGTGGCGGGTCCCTGAGCTCCCTAGAGTCGGCCACACAGATTGAGACTTGGACTATGA
 TTATCTACAGAACTGGGACCTCGTTAAGAAACTAGCAGATTGTATGGTTCAAAGACACTTTGATGACGATT
CTTAACAATAACGATAAAATTGGCCTTAAGAACTGTGCTGGCGTCTCAAGAACATAGAAGATGTGTAACAGGT
 ATTTTTT

FIGURE 2

AACTCAAACCTCTCTGGGAAAACGCGGTCTGCTCCTCCGGAGTGGCCTGGCAGGGTGTGGAGCCCTCG
GTCTGCCCGTCCGGTCTCTGGGCCAAGGCTGGTTCCCTCATGTTATGGCAAGAGCTACTCGTGC GG TGCTTC
TTCTCCTGGCATACAGCTCACAGCTTTGGCTATAGCAGCTGTGAAATTATACCTCCGGTGCTGGAGGCT
GTTAATGGGACAGATGCTCGTTAAATGCACTTCTCCAGCTTGCCCTGTGGGTGATGCTCTAACAGTGACCTG
GAATTTCGTCCTCTAGACGGGGGACCTGAGCAGTTGATTCACACCACATAGATCCCTCCAACCCATGAGTG
GGCGGTTAAAGGACCGGGTCTTGGGATGGGAATCCTGAGCGGTACGATGCCTCCATCCTCTTGAAACTGCAG
TTCGACGACAATGGGACATACACCTGCCAGGTGAAGAACCCACCTGATGTTGATGGGTGATAGGGAGATCCGGCT
CAGCGTCGTGCACACTGTACGCTCTGAGATCCACTTCCTGGCTCGGCCATTGGCTTGCACTGATGTA
TCATAATAGTAATTGTTAGTGGCCTCTCCAGCATTACCGAAAAAGCGATGGGCCAAAGAGCTCATAAAGTGGTG
GAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCTGTTATTTAGAAGACACAGACTAAACA
ATTTAGATGGAAGCTGAGATGATTCCAAGAACAAAGAACCCCTAGTATTCTGAAGTTAATGGAAACTTTCTTG
GCTTTCCAGTTGTGACCCGTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGC
CAGCACAGTGCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTATTAAATTTCAGAGTGTAAATT
TTTCAAGTGCTCATTAGTTTATAAACAAAGAAGCTACATTGCCCCCTAAGACACTACTTACAGTGTATGACT
TGTATACACATATGGTATCAAAGGGATAAAAGCCAATTGCTGTTACATTCCCTTCACGTATTCTTTAG
CAGCACTCTGCTACTAAAGTTAATGTGTTACTCTCTCCACATTCTCAATTAAAGGTGAGCTAACGCT
CCTCGGTGTTCTGATTAACAGTAAATCTAAACTGTTAAATGACATTTTATTTATGTCTCCTTAA
CTATGAGACACATTTGTTACTGAATTCTTCAATATTCCAGGTGATAGATTGTCG

FIGURE 3

GGCACGAGGCGATTCAAGGGGAGGGAGCAACTGGAGCCTCAGGCCCTCAGAGTAGTCTGCCTGACCACCCCTGGAGCC
CACAGAAGCCCAGGACGTCTCCCGCGAGGCCTCCCGTGTGGCTGAGGATGGCTGAGCAGCAGGGCCGGAGCTT
GAGGCTGAGTGCCCCGTCTGCTGGAACCCCTCAACAACACGTTCCATACCCCCAAAATGCTGGATTGCTGCCACTC
CTTCTGCGTGGAAATGTCTGGCCACCTCAGCCTGTGACTCCAGGCCGGCGCCGCTGCTGTGCCACTCTGCGCC
AGCCCACAGTGTGGCCTCAGGGCAGCCTGTCACTGACTTCCCACGGACACTGCCATGCTCACCCCTGCTCCGCC
GAGCCCCACCATGTCACTCTGGAAGGCCATCAGCTGTGCCCTCAAGGACCAGCCAAAGAGCCGCTACTTCCTGCC
GCCTCGAGTCTACACGCTGGACCTTGGCCCCCAGCCTGGGGCCAGACTGGGCCGCCCCAGACACGCCCTTGCCA
CCGTGTCTACGCCATCCTCATCCCCAGCCACACTTTGAGGGAGTGTTCGCCAACCTCAGTTCCGATCTT
GCCTACCTGATGGCGTCATCCTCAGTGTCACTCTGCTCATATTCTCCATCTTGGACCAAGCAGTCCTT
GGGTGTGGGGTGAGTGTCTTCCCAGACAAGAAACAAACCTTTGGTTGCTGCTGGGTATGGTACTACGGAGC
CTCATTGGTATTGTCTCCTTGTAGTGTGTTATTTACAATCCAGGGATTGTTAGGCCATGTGTTGCTTCT
GGGAACAATTAAAAAAAAACAAAAACGAAAAGCTTGAAGGACTGGGAGATGTGGAGCGACCTCCGGGTGTGAG
TGTGGCGTCATGGAAGGGCAGAGAAGCGGTTCTGACCACAGAGCTCCACAGCAAGTTGTGCCAAAGGGCTGCACAGT
GGTATCCAGGAACCTGACTAGCCAAATAGCAAGTTGCATTCTCACTGGAGCTGCTCAAAATCAGTCATATT
TTTGAGTTGCTCTTACTATGGGTTGCTAAAAAAAAAAATTGGGAAGTGAGCTTCAATTCTGTGGGTA
AATGTGTGTTGTTCTTTGAATGTCTTGCCTGGTTGCAGTAAAAGTGTCTGTATTCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

GAAGCGCGCTCCGGGGAGGTGTTGCAGCCATGGCTACGGCAGCCGGCGACCTACTTCAGCGAGGCAGTCTGTT
 CTGGTTACAGTCATCACCCCTCAGCTTGGCTACTACACATGGGTTGCTCTGGCCTCAGAGTATCCCTATCAGA
 ACCTTGGGCCCTGGGCCCTCACTCAGTACTTGGTGGACCACCATCACACCCCTTGCATGGGTATTGGCTT
 GCCTGGCTGATTCATGTGGAGAGTCCTGTATGCCATAGCATTGTGCAAGCATAAAGGCATCACAAGTGGTCGGC
 TCAGCTACTCTGGTCCTACAGACTTCTTGGATAGCGTCTCACCATTTGATTGCTTACAAACGGAAGC
 GCCAAAAACAAACTTGAAGTTGTCTGAAAGCTTGCTCTACACTTACATTCATCCTCACCCCTTTTGTGGGT
 AGAGGAGGTGCAGTAATTACTCAGTGATCTTCTACTTCTAGAAACTGTCCTCAAAGCTTTAAGACCCCCCTC
 GTTAGTCAGTTCTCTTTATATGCTCTGGTGAAGTACAGCTTGTACTTAAGAAAGAACAGAGAAA
 GATTTAGCTTCAATCCTATTGGCAGAGGACTTCAGCTACCTTCTACAGTCTTGGCTGTGTTGGTACCCCTCG
 TGTGCTCTGAGCTAACGCCACATACTAAACTGACTTTGGTTGTATACCCCTGCTCCGCCTCTGATGAAAACAC
 CTTACCCCTACAACCACCATTTCTCCTTCAAAGCTTTCCACCTTGCTGCACTAAGATAAAGTGA
 ACTTCCACTATATGTCAATTCCACACATTTATTAGGTACCTGTGAGGTAGGATCCTATCCTCTCAAACCTCATTCT
 CATGCTACAGAGAAAGATAAGGAAGATGAGCAAGTGCCTGGAATGGGCAGGCTGAGCAGTCACACAGGCATAGAGG
 CACGCTGAGAACCTGGAGGGGAGACTGCAGAGTGCCTCCCTGATGCTGCAGCCGGAAAGTGATCCTCCACCT
 GGCCCTGGGACACTGTGCTCTGCAGTGTGCAGGGCCTGATGGCACTGCTAGATTGCTCCTCAGCTCAGGGCCACA
 GCTTAAACAGCTTACCTTCCCTCAGCACCTGCTCCACTATCTGCACACAGGTGCTCTAACCATGTTATTGAA
 CAAAGGAGGGAAACTGATTTCACTTCACTTGTTCATTATCATTCAAATTGTTATGTGAAAATGGCACAACCCATT
 GGGTACCCCTCACCCAAAATAAAAGCCAAGTCTACCTTGACTGGTACCACTTTGTGGTTCTGGTGA
 AACCTTATCTTCTACCTTCTATTCTCAATCACTCTCCAAAAGTGTGCTTCCAGCTCTGATTTATTCA
 AACACAAGCATTCTGTTAGAGATTCTAGCCCATGGGTATCTGGCTAGTTATTACCTCTCCTGTTCACTTAGTT
 ATACTTATTATTGCTCACAGGCTGGGAGGCAGAATGACTCTGTCACCACTAGGAGCCATTAGGGCTTCTCCCTG
 GAGGACTGCCTGCTTGTGCTGGGACACTAGCCCTCATTCCCTCTGTGGTACAGTGGGCAAATTATTGTAT
 TAAGCAAACATTATGGAAACAACCCGCTCCGAAAACGGAGCCCCAAGTAAAGCACAACCCCTGAAAGATTATGA
 ACTATGAATTGTCTCTAGTAGAGATAAATTCTGCAAACATATCTCAGTCTCCCTGTGTTCTGGTATTAAGA
 AGTTCCCTTGGTAAGGAAAAGGATTTTAACCATAGAGTTAGGCATCATGAAATTCAAACCAAGATTCTTAATA
 CCTGGTCTCCTCAAAGAGAAATAAACAGTAATAGGGTGGCTGGGAACAATATGGCAGATTATTGAATGAAATTG
 ATTAACCTGAATAAAATGCTGTGAATTTC

FIGURE 5

GGCACGAGGCCGAGCGGACTGCCCTTCCCAAGATGGCGTCGAAGATAGGTTCGAGACGGTGGATGTTGCAGCTGA
 TCATGCAGTTGGGTCGGTGCACACGCTGCCCTTGGGCTGCTCAGCCAGCTCATGCTGTACGCTGAG
 AGGGCTGAGGCACGCCGAAGCCGACATCCCAGTGCCTACCTGTATTCGACATGGGGCAGCCGTGCTGTGCGC
 TAGTTCATGTCCTTGGCGTAAGCGCGCTGGTCGCGCTGGGGCCGACTCCAATTGCCATTAGCACCTACG
 CCGCCTACATCGGGGCTACGTCCACTACGGGACTGGCTGAAGGTCCGTATGTACTCGCGCACAGTGCACATC
 GCGGCTTCTTGTGTTGGCCAGCGGTGCTGGGAGCTGTACGCCGAAACCTCGCAGCCGCTCCCTGCAGTCAC
 CGGCCAGGTGTTCTGGTATCTACCTCATCTGTGTCGCTACTCACAGCACAGCAAGGAGGACCGGCTGGCGT
 ATCTGAACCATCTCCAGGAGGGAGCTGATGATCCAGCTGTTCTCGTGTATGGCATCCTGGCCCTGGCC
 CTGTCAGGCTACTACGTGACCCCTCGCTGCCAGATCCTGGCTGTACTGCTGCCCTGTATGCTGCTCATTGATGG
 CAATGTTGCTTACTGGCACAAACACGCGCGTGTGAGTTGGAACCAGATGAAGCTCCTGGAGAGAGTGTGGCA
 TCTCGGAACTGCTGTCATCCTGCCACTGATGGCTGAGTTTATGGCAAGAGGCTGAGATGGCACAGGGAGCCAC
 TGAGGGTCACCTGCCTCCCTGCTGGCCAGCTGCTGTTATTATGCTTTGGTCTGTTGATCTT
 TGCTTTTAAATTGTTTGCAGTTAAGAGGCAGCTCATTGTCAAATTCTGGCTCAGCGCTGGAGGGC
 AGGAGCCCTGGCACTAATGCTGTACAGGTTTCTGTTAGGAGAGCTGAGGCCAGCTGCCACTGAGTCCTG
 TCCCTGAGAAGGGAGTGGCAGGGCTGGGATGCGGCTACTGAGAGTGGGAGAGTGGGAGACAGAGGAAGGAAGATG
 GAGATTGGAAGTGAGCAAATGTGAAAAATTCTCTTGAACCTGGCAGATGCAGCTAGGCTCTGCAGTGCTGTTGG
 AGACTGTGAGAGGGAGTGTGTGTTGACACATGTGGATCAGGCCAGGAAGGGCACAGGGCTGAGCACTACAGAA
 GTCACATGGTTCTCAGGGTATGCCAGGGCAGAAAACAGTACCGGCTCTGTCACTCACCTTGAGAGTAGAGCAGA
 CCCTGTTCTGCTGGCTGTGAAGGGTGGAGCAGGCAGTGGCCAGCTTGCCCTCCTGCTGCTCTGTTCTAG
 CTCCATGGTGGCCTGGTGGGGGTGGAGTCCCTCCAAACACCAGACCACACAGTCCTCCAAAATAAACATTAA
 TATAGACAAAAAAAAAAAAAAA

FIGURE 6

GGCCGCGGCTGCCCTGGCCCTTATCAGGATGAAAACGCTCTGTTGGTGTCTGGGCCCTGCTGGCCTTGAT
 CCTTGCCCAGGGTCCCGGAAGAGTTGAGGTTCTATTGCCAAGTCAGGCCCTGGTGGAGTTGGACAGT
 CCCTAGTGTGCAACTGCAGCACTACTGCCAGACCCAGGACCCAGTGAATTGAGACCTTCTAAAGAAAACCTCAG
 GTGGACAAAGGGCCTCAGTGAAGAGAGTTCTCTGGAGGATGTCACAGAGAATTCCATCCTGCAGTGCTTCTTC
 TTGTGCAGGGATTCAAAGGACACAAGCCTTGGCATCACTGTGTATCAGCCACCAGAGCAAGTGATCCTGGAGCTGC
 AGCCTGCCCTGGGTGGCGTGGACGAAGCCTCACAGTGAAGTGTATGTACCCAGTGTAGCACCCCTGGAGAGTCTC
 ACCCTGCCCTCTCCAGGGTAACCAAGAACTGCATAGAAAGAACTTACGAGCTGGCTGGCCTCCAAAGAGC
 TGAAGTCATCATCAGTGTCAAGGCCAAAGGAGAATGACAGATGCAATTCTCCTGCCATGCAGAACTGGACTTGA
 GTTGCAAGGTGGGAGGCTTTCAAGGCAGCTACCCATCAGAATAGTCGGATCTTGAAATTCTCAGAGTCCC
 CACATCTGGGTCTCTCCCTTGGAGGCTGGATGGCGGAGACTGTGAGCTGCGAGGTGGCTAGGGTGTTCAGC
 CAAAGAAGTTATGTTCCACATGTTCTGGAAAGACCAAGAGCTGAGCTCCCTCCCTGGAGGGGACACAGCAT
 GGGCAATGCTACCATTGGACCATGGAGGCTGGTGTAGGAACTGTCTTGATCTGGTCAATGGAA
 CAGAAGACAAGAAAGCTAGTCATAGCTACAGCTCCCTCCACCAATCCTGGAGCTAAAGAATCATACCCATTGGC
 AGGGACCGACATTAATGTGACCTGCTCAGGGCATGTATTAACATCACCAGCCCTACTCTCGGCTCAGGGAGCCC
 CAGACCTCCCTGCTGGGAGCCTGCCCTACTACTGCCAGGGAGGAAGATGATGCCXGAAATTCTCCTGC
 GAGGCCTTTGGTGGTGCAGGGTCAGCGGTTGATGAAAACCACTGTGATCCAGCTCCATATCCTAAAGCCACAGTT
 AGAGGAATCCAGTTGCCCTGGCAAACAGACCTGGCTGGAAGGGATGGAACACACGCTGCCCTCGTCCAAAGGGAA
 ACCCAGCTCCAGCCTGGTGTACCTGGAATGGGTGGTCTTGACCTTGAAGTGCCACAGAAGGCAACCTAGAAC
 CACACTGGAACCTACCGCTACACAGCCACTAACAGCTGGCTCTGTCAGCAAAGACATTGCTGT

FIGURE 7

CCACCGCGTCCGTTCTGAGGTGCATTCTTTTGATGAGAGGCATCTCTAGGTACCATCCCTGACCTGGTCCT**CATG**
 CTGCCGAGGCTGTTGATCTGTGCTCCACTCTGTGAACCTGCCAGCTGTTTGATGCCAGCCCCCTCCA
 TCCCACAGAGGGGAGCCAGTGACCTGACGTGTAAGATGCCCTTCTACAGAGTTAGCAGATGCCAGTTCCAGTTCT
 GCTTTTCAGAGACACCCGGCCTGGGCCAGGCTGGAGCAGCTCCCCAAGCTCCAGATCGCTGCCATGTGGAAA
 GAAGACACAGGGTCATACTGGTGCAGGCACAGACAATGGCTCAAAGTCTTGAGGAGCAGGAGATCCCAGATAAA
 TGTGCACATCCCGGTGTCGCCAATCCTCATGCTCAGGGCTCCCAGGGCCAGGCTGCAGTGGAGGATGTGCTGG
 AGCTTCACTGTGAGGCCCTGAGAGGCTCTCCTCCAATCCTGACTGGTTTATCAGGAGATATCACCCTGGGGAGC
 AGGTCGGCCCCCTCTGGAGGAGGAGCCTCCTCAACCTTCCCTGACTGAAGAACATTCTGGAAACTACTCCTGTGA
 GGCAACAATGGCTGGGGCCCAGCGCAGTGAGGCGGTGACACTCAACTCACAGTGCCTACTGGGCCAGAACGA
 ATCATCTTACCTCAGGAGTCATTGAGGGCTGCTCAGCACCCCTGGTCCAGCCACCGTGGCCTTATTATTTGCTAC
 GGCCTAAAAGAAAAATAGGAAGACGTTCAGCCAGGGATCCACTCAGGAGCCTCCAGCCTAACCCAAAGAGTTCAC
 CTACCTCAACTCACCTACCCAGGGCAGCTACAGCCTATATATGAAAATGTGAATGTTGTAAGTGGGATGAGGTT
 ATTCACTGGCGTACTATAACCAGCCGGAGCAGGAATCAGTAGCAGCAGAAACCCCTGGGGACACATATGGAGGACAAG
 GTTCCCTAGACATCTATTCCAGGCTGAGGAAGCAAACATTACAGATGTGGACTATGAAGATGCTAT**GTAAAGTTA**
 TGGAAAGATTCTGCTTTGAAAACCATCCATGACCCAAGCCTCAGGCCCTGATATGTTCTCAGAGATCCTGGGCA
 TTAGCTTCCAGTATACTCTGGATGCCATTCTCCATGGCACTATTCTTCACTGTGAAGTGAAGTGGC
 GCAGCCCTGAAGAAACTACCTAGGAGAACTAATAGACACAGGAGTGACAGGGACTTGTATCAGAACAGATTCC
 GCCGGCTCCTTGAAAACAGGTATATTGTCCTCTGTTACAAGAGGAAACAAGATGGAATAAAAGAAATTGGG
 ATCTTGGGTTGGAGGGACAGTGAAAGCTAGAGCACATGAACACTAAGGTTAGTGAATCTGCAGGACTTCACAGAGAGA
 GCTGTGCCCATCTCAGTCCAAGTGCTTCTGCCAGACAGCACAGAACTCCAGCCCCGCTACTTACATGGATC
 ATCGAGTTCCACCTAAAATATGATTCTATTGAGTCAGTGTACCTGTTACCAATTAGAACACTAAAACAAAGTTACAT
 AAAAGTTATTGTGACTCCACTTAATTAGTGAAGCTTATAAAATACAGTGTGTCCTCTTTATTCAACAAATTGGAAATCGTG
 TATGTATCTATTACAGCCCTAGAACGCTTATAAAATACAGTGTGTCCTCTTTATTCAACAAATTGGAAATCGTG
 GTAATATGGTTGAAACCTGTATCTTAATTATTTTTAAATTGAGACAGGGCTCAGTGTCAACTCAATCTG
 GAATGCAGTGGCACAATCTGCCACTGCAACGCCCTGCCAGGCTCAAGCAAACCTCTCACCTCAGCCTGCTG
 AGTAGCTGGACTACAGGCACATGCCACAAACTGGCATTGGCTTACGTAGAGAACAGATTCAACCGTT
 GCCCAGGCTGGCTCAAACCTCTGGCTCAAGCAATGTATTGAATTAAACCAAGGCACTCACTCTTATGAAT
 TAATAAACATTGGAGGTATATAAGTAAAAGTAAAGTCTTCTGTAAGTTAACACAAATGTTAACTATTGTTA
 AAAACTTACAGGTAGCTCTAGATATTCTATTGTATGTACTTATGCATACATGTAAGTATATAAAC
 TTAGAAGTGTACCTATCTAACAAACTATTATGAAATACTTCAAACTGTAAATAGATCTATTACTATTAA
 AGTCTCTATAGTAGTGTGTTATAGATAATCATAACTTTCTTTTATTGAGTAAATATGCACAAACATAA
 AATTGATCATTAAACCATTAAAGTGTACAATTCACTGAGGACTCTCAAAAAATTCAATTATAAACTTAAATCCTCTC
 ATCACTGGTGGACATTAAGGAGACTCTCAAAAAATTCAATTATAAAACAAAGTCAAAACAAATGTCTTGTACTA
 GCATATTATGCCACTCTGCTGGATTATCTGAAGGATAAATTGTAATCTAGTATTGCTAGATTATGCATATTAAA
 TATTCTGTTAAATAGTCAAAAAAAAAAAAAA

FIGURE 8

CTCAATCAGCTTATGCAGAGAAGAAGCTTACTGAGCTCACTGCTGGTGCTGGTAGGCAAGTGCTGCTTGGCAA
 TCTGGGCTGACCTGGCTCTCCTCAGAACTCCTCTCCAACCCCTGGAGCAGGCTTCAATGCTGCTGCTGGCGTCC
 TTGCTGCCCTTGCTCCAGTCTGTGGACAATCTGCAGCTGCACACAAACCTGTGATTCCTGAGGATTCAGTCTATGCAACAGAGAAAACAACATGGT
 CACATTCTCAAAGGAGAGAGAGTGACTCTGACTGCAATGGATTCAGTCTATGCAACAGAGAAAACAACATGGT
 ATCATCGGCACTACTGGGAGAAAAGTGACCCCTGACCCAGGAAACACCCCTCGAGGTCGGGAATCTGGACTGTAC
 AGATGCCAGGCCGGGCTCCCACGAAGTAACCCCTGTGCGCTTGCTCTTCAGACTCCTAATCCTGCAGGC
 ACCATATTCTGTGTTGAAGGTGACACATTGGTCTGAGATGCCACAGAAGAAGGAAAGAGAAATTGACTGCTGTGA
 AATATACTTGGAAATGGAACATTCTTCCATTCTAATAAAAGCTGGATCTTCTATCCCACAAGCAAGTTCAAAT
 ACAATGGCAATTATCGATGCATTGGATATGGAGATGAGAATGATGTATTAGATCAAATTCAAATAATTAAAAT
 TCAAGAACTATTCCACATCAGAGCTGAAAGCTACAGACTCTCAGCCTACAGAGGGGAATTCTGTAACACTGAGCT
 GTGAAACACAGCTCCTCCAGAGCGGTACAGACACCCCACTTCACCTCAGAGATGGCGAGGTACATCCTG
 TCAGACTGGAGCACGTACCCCGAACCTCAGCTCCAACCGTCTGGAGAGAAAACCTCAGGATCCTATTGGTGTGGC
 TGAAACAGTGGGGTAACATCCACAGCACAGTCCCTCGTACAGATCCATGTGCGAGGGATCCCTGTGCTGGG
 TGCTCCGGAGACCCAGCCCTCAGGGGGCAGGCTGTGAGGGGGAGATGCTGGCCTTGTGCTGCTCCGGCTGAA
 GGCACAGGGGATACCACATTCTCCTGGCACCATCAGACAGAGCCATGCAGGGGGATACTACTGTACAGCAGACAACAGCTACG
 GCCCTGTCCAGAGCATGGTGTGAATGTCACTGTGAGAGAGACCCAGGCAACAGAGATGGCCTTGTGCGCCGGGA
 GCCACTGGAGGGCTGCTCAGTGTCTTCTCCTGGCTGTGGCCCTGCTGTTCACTGCTGGCGTGGAGGAAGTCAGG
 AGTTGGTTCTGGAGACGAAACCAGGCTCCCTCCCGTCCAGGCCCAGGAGAGTCTCCATTCCATCTGCCCTG
 CCCAGGGAGCTCAGTGTGATGTACACCCAAAAGGGAGATTGGTATACTCTGAGATCCAGACT
 ACTCAGCTGGAGAAGAGAGGAAGCTAATACCTCAGGACACTCTAGAGGATAAGGATGTCAGTGTACTC
 TGAGGTAAAGACACAACACCCAGATAACTCAGCTGGAAAGATCAGCTAAGGATGAAGAAAGTTAAAGAGAATGAAA
 AGTTACGGGAACGTCCTACTCATGTGATTCTCCCTGTCAAAGTCCCAGGCCCAGTGCAGTCCTGCGGCACCTG
 GAATGATCAACTCATTCCAGTTCTAATTCTCATGCATATGCATTCACTCCAGGAATACTCATTCTGCTACT
 CTGATGTTGGGATGGAATGGCCTGAAAGACTCACTAAATGACCAGGATCCACAGTTAAGAGAAGACCCGTAG
 TATTTGCTGTGGGCTGACCTAATGCATTCCCTAGGTCTGCTTTAGAGAAGGGGATAAAAGAGAGAGAAGGACTGT
 TATGAAAAACAGAACAAATTGGTGAATTGGGATTTGCAAGAGATGAAAAAGACTGGGTGACCTGGATCTCTGC
 TTAATACATCTACAACCATTGTCCTACTGGAGACTCACTGCACTCAGTTGTTAACTGTGAGTGGCTGCACAGGC
 CTGCAACAAATGAAAAGCCCTTCACCTCTGCTGCACAGCTACACTGTCAAGGATTGCAAGGATTTAAAGAA
 CCCATCTGGAATGGTTACAGAGAGAGGAATTAAAAGAGGACATCAGAAGAGCTGGAGATGCAAGCTCTAGGTGC
 GCTTCCAAAAGCAAATGATAATTGTAATGTCATTAGTGCACAAAGATTGCAACATTAGAGAAAAGAGACACAAA
 TATAAAATTTAAACTTAAGTACCAACTCTCCAAAATTAATTGAACTTAAATATTGATTAACACTCATAATAAAA
 CTCTGCCCTTAAAAAAAGATAAAATTCTCCTACGCTGTCACTGAAATAATTACCAACCCCTAGCAATAAGC
 CTTGCAAGAGAGGTTTATTCTCTAAATACCATTCCCTCTCAAAGGAAATAAGGTTGCTTTCTGTAGGAACGT
 GTCTTGAGTTACTAATTGTTATATGAGAATAATTGCAATAAAATGAGAAGGAATAAAAGAAATAGGAAGCC
 ACAAAATTGATGGATATTGATGACACCTACTGGTAAATAATTGACAAAAACAGCAGGCCAAATATTAGAGGT
 CTCCGTGATGGAAGTGTACAATACCACCTACAAATTATCCATGCCCAAGTGTAAAAGTGAATCCATTCAAGTCTT
 CTAACTGAATACTGTTTATAGAAAATGCATGGAGAAAAGGAATTGTTAAATAACATTATGGGATTGCAACCC
 CAAAACATAAAACTGAGAAAAAGTTCTATAGGGCAAATCACCTGGCTTCTATAACAAATAATGGGAAAAAAATGAAA
 TAAAAAGAAGAGAGGGAGGAAGAAAGGGAGAGAGAAGAAAAGAAAATGAAGAAAAGTAATTAGAATATTTCACAA
 TAAAGAAAAGACGAATATTAAAGGTGACAGATATCCCAACTACGCTGATTGATTTACAAATTATGAGTGTAT
 GAATTGTCACATGTATCACCCCAAAAAAGAGAAAAGAAAATAGAAGACATATAAAATTAAATGAGACGAGACA
 TGTCGACCAAAAGGAATGTGGGTCTGTTGGATCTGACTCAAATTAGAAAAAAACTACCTACGAAATA
 CTAAGAAAATTGTTACTAATATTAAAGAAATTGTTGTTGGATATAAGTGTAGTTATTGAGTGT
 TTTTATAAAAGCAAAAGGATATTCACTTCAGCGCTTACACTGAAAGTATTAGATTAAAGCTTATTAACGTA

FIGURE 9

GCCGAGCTGAGCGGATCCTCACATGACTGTGATCCGATTCTTCCAGCGGTTCTGCAACCAAGCGGGTCTTACCC
 CGGTCCCTCCCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGAGAGTCCCCGAATCCCCGCTCCAGGCTA
 CCTAAGAGGATGAGCGGTGCTCGACGGCGGGGAGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGC
 TCAGGGCGGACCGTGAGTCCAAGTCGCCGCCTTGCCTGGAGAGATGAATGTCCTGGCGACGGACTCC
 TGCAGCTCGGCCAGGGCTGCGGAACACGCGGAGCGCACCCGAGTCAGCTAGCTAGCGCGCTGGAGCGGCCCTGAGC
 GCGTGGGGTCCGCTGTCAAGGAACGAGGGTCCACCGACCTCCCGTAGCCCTGAGAGGCCGGTGGACCCCTGA
 GGTCCCTCACAGCCTGCAGACACAACCTCAAGGCTCAGAACACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCAGC
 AGCAGCGGCACCTGGAGAACGAGCACCTGCGAATTAGCATCTGCAAAGCCAGTTGGCCTGGACCACAAGCAC
 CTAGACCATGAGGTGGCAAGCCTGCCGAAGAAAGAGGCTGCCAGATGGCCAGTTGACCCGGCTCACAA
 TGTCAGCCGCCTGCACCGGCTGCCAGGGATTGCCAGGAGCTTCCAGGTTGGGAGAGGCAGAGTGGACTATTG
 AAATCCAGCCTCAGGGTCTCCGCATTTGGTAAGTGCAGATGACCTCAGATGGAGGCTGGACAGTAATTAG
 AGGCGCCACGATGGCTCAGTGGACTTCAACCGGCCCTGGAAAGCCTACAAGGCGGGTTGGGATCCCCACGGCA
 GTTCTGGCTGGGCTGGAGAACGGTGCATAGCATACGGGGGACCGAACAGCCGCCTGGCGTGCAGCTGGGGACT
 GGGATGGCAACGCCAGTTGCTGCAGTTCTCGTGCACCTGGTGGCAGGACACGGCTATAGCCTGCAGCTCACT
 GCACCCGTGGCGGCCAGCTGGCGCCACCACCGTCCCACCCAGCGGCCCTCCGTACCCCTCTCCACTGGGACCA
 GGATCACGACCTCCGCAGGGACAAGAAACTGCGCCAAGAGCCTCTGGAGGCTGGTTGGCACCTGCAGCCATT
 CCAACCTCAACGCCAGTACTCCGCTCCATCCCACAGCAGCGGCAGAACGCTTAAGAAGGAAATCTCTGGAAGACC
 TGGCGGGCCGCTACTACCCGCTGCAGGCCACCACATGTTGATCCAGCCCAGGCAGCAGAGGCAGGCTCCTAGCG
 TCCTGGCTGGCCTGGTCCCAGGCCACGAAAGACGGTGACTGGCTCTGCCAGGGATGTGGCGTCCCTGCC
 TGGCAGGGCTCCAAGGAGGGCAGTGGAAACTTGTGGACAGAGAACGACCACGACTGGAGAACGCCCTTT
 CTGAGTGCAGGGGGCTGCATGCCTGCCCTGAGATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCGTGGAC
 CAAGGGCATGGAGCTTCACTCCTGCTGGCAGGGAGTTGGGACTCAGAGGGACCACTGGGCCAGCCAGACTG
 GCCTCAATGGCGACTCAGTCACATTGACTGACAGGGACCAAGGGCTTGTGGTGGCGAGAGCGCCCTCATGGTGTG
 GTGCTGTTGTGTAGGTCCCTGGGACACAAGCAGGCCAATGGTATCTGGCGGAGCTCACAGAGTTCTTGG
 ATAAAAGCAACCTCAGAACAC

FIGURE 10

GATGTGCTCCTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTGGAATTGAGGAAACT
TCTCTTGATCTCAGCCCTGGTGGTCCAGGTCTTCATGCTGCTGTGGTGATATTACTGGTCCTGGCTCTGTCA
GTGGACAGTTGCAAGGACACCCAGGCCATTATTTCCCTCAGCCTCCATGGACCACAGTCTCCAAGGAGAGAGA
GTGACCCCTCACTTGCAAGGGATTCGCTTCATCTCACCACAGAAAAGGTACCCGTACTGGAAAGA
AATACTAAGAGAAACCCCAGACAATATCCTTGAGGTTCAGGAACTTGGAGAGTACAGATGCCAGGCCAGGGCTCC
CTCTCAGTAGCCCTGTGCACTTGGATTTTCCAGAGGATGGGATTCTCATGCTGCCCAGGCTAAGTGAACTC
CTGGGCTCAAGTGATCTGCTCACCTAGGCCCTCAAAGCGCTGGGATTACGCTCGGCTGATCCTGCAAGCTCCACT
TTCTGTGTTGAAGGAGACTCTGTGGTTCTGAGGTGCCGGCAAAGCGGAAGTAACACTGAATAATACTATTACA
AGAATGATAATGTCCTGGCATTCTTAATAAAAACTGACTTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 11A

CGCCCGCGCGCTGCAGCCCCATCTCCTAGCGGCAGCCCAGGGCGGGAGCGAGTCGCCCGGAGGTAGGTCCAG
 GACGGGGCGCACAGCAGCGAGGCTGGCCGGGAGAGGGAGGAAGAGGGATGGCAGGGCCACGCCCGAGCCATGGG
 CCAGGCTGCTCTGGCAGCCTTGATCAGCGTCAGCCTCTCTGGGACCTGGCAAACCGCTGCAAGAAGGCCCCAGTG
 AAGAGCTGCACGGAGTGTGTCGTGGATAAGGACTGCGCCTACTGCACAGACAGAGATGTTAGGGACCGGGCCTG
 CAACACCCAGGCGGAGCTGCTGGCCGGGCTGCCAGCGGGAGAGCATCGTGGTCATGGAGAGCAGCTTCAAATCA
 CAGAGGAGACCCAGATTGACACCACCCCTGCGGCGCAGCCAGATGTCCCCCAAGGCCTGCGGGTCCGTCGCGGCC
 GGTGAGGAGCGGCATTTGAGCTGGAGGTGTTGAGCCACTGGAGAGCCCCGTGGACCTGTACATCCTCATGGACTT
 CTCCAACCTCATGTCCGATGATCTGGACAACCTCAAGAAGATGGGGCAGAACCTGGCTCGGGTCCCTGAGCCAGCTA
 CCAGCGACTACACTATTGGATTGGCAAGTTGTGGACAAAGTCAGCGTCCCGCAGACGGACATGAGGCCTGAGAAG
 CTGAAGGAGCCCTGGCCAACAGTGACCCCCCTTCCTCAAGAACGTCACTCAGCCTGACAGAACAGATGTGGATGA
 GTTCCGGAATAAAACTGCAGGGAGAGCGGATCTCAGGCAACCTGGATGTCCTGAGGGCGGCTTCGATGCCATCCTGC
 AGACAGCTGTGTCGACGAGGACATTGGCTGGCCGGACAGCACCCACCTGCTGGTCTTCCTCACCAGACTCAGCC
 TTCCACTATGAGGCTGATGGGCCAACAGTGTGGCATCATGAGCCGAAAGATGAACGGTGCACCTGGCACCTGGACAC
 CACGGGACCTACACCCAGTACAGGACACAGGACTACCCGCGGCGACAGCACCCACCTGGCTGCGCTGCTGCCAAGCACA
 ACATCATCCCCATTTGCTGTCACCAACTCTCTATAGCTACTACAGAGAACGCTTCAACACCTATTTCCCTGTCCTCC
 TCACTGGGGTGTGTCAGGAGGACTCGCCAACATCGTGGAGCTGCTGGAGGAGGCCCTCAATCGGATCCGCTCCAA
 CCTGGACATCCGGCCCTAGACAGCCCCGAGGCCCTCGGACAGAGGTACCTCAAGATGTTCCAGAACAGAGGAGGA
 CTGGGTCTTTCACATCCGGGGGGAGTGGTATATACCAGGTGAGCTGCGGGCCCTTGAGCACGTGGATGGG
 ACGCACGTGTGCCAGCTGCCGGAGGACAGAACGGCAACATCCATCTGAAACCTTCCTCTCCGACGCCCTCAAGAT
 GGACGCGGGCATCTGTGATGTGTCACCTGGAGCTGCAAAAGAGGTGCGGTGAGCTGCGCTGAGCTTCAGCTCAACG
 GAGACTTCGTGTGCCAGTGTGTGTCAGCGAGGGCTGGAGTGGCCAGACCTGCAACTGCTCCACCGGCTCTG
 AGTGACATTCCGGCTGCTGCCGGAGGGCGAGGACAAGCCGTGCTCCGGCGTGGGAGTGGCAGTGCAGGGACTG
 TGTGTGCTACGGCGAAGGCCCTACGAGGGTCAGTTCTGCGAGTATGACAACATTCCAGTGTCCCCGACTTCCGGT
 TCCTCTGCAATGACCGGAGGACGCTGCTCCATGGGCCAGTGTGTGAGCCTGGTGGACAGGCCAAGCTGTGAC
 TGTCCCCCTCAGCAATGCCACCTGCATCGACAGCAATGGGGCATCTGTAATGGACGTGGCACTGTGAGTGTGGCG
 CTGCCACTGCCACCGCAGTCGCTACACGGACACCATCGCAGATCAACTACTCGGCGATCCACCCGGGCTCT
 GCGAGGACCTACGCTCTGCGTGCAGTGCCAGGGTGGGAGAAGAACGGGCGCACGTGTGAGGAATGC
 AACTTCAAGGTCAAGATGGTGGACGAGCTTAAGAGAGCCGAGGAGGTGGTGGCTGCTCTCCGGGACAGGAGGA
 TGACGACTGCACCTACAGCTACCATGGAAGGTGACGCCCTGGGGCCACAGCACTGTCTGGTGGCACAAGA
 AGAAGGACTGCCCTGGGCTCTCTGGCTCATCCCCCTGCTCTCCCTGGGCTCTGGGGCTCTGGGCTCAT
 CTGCTGCTATGCTGGAAAGTACTGTGCTGCAAGGCCCTGCTGGCATTCTCCCGTGTGCAACCGAGGTACAT
 GGTGGGTTTAAGGAAGACCAACTACATGCTGCCGGAGAACCTGATGGCTCTGACCAACTTGGACACGCCATGCTG
 GCAGCGGGAACCTCAAGGGCGTGCAGTGGCCCTGGAAAGGTACCAACAACTGCAAGGCCCTGGGTTGGC
 CATGCCGCCAGCATCAACCCACAGAGCTGGTGCCCTACGGGCTGTCTGGCCTGGGCCCTTGCACCGAGAA
 CCTGCTGAAGCCTGACACTCGGGAGTGCGCCAGCTGCCAGGGAGTGGGAGGAACCTGAACGAGGTACAGGC
 AGATCTCCGGTGTACACAAGCTCAGCAGACCAAGTCCGGCAGCAGGCCAATGCCGGAAAAAGCAAGACCACACC
 ATTGTGGACACAGTGTGATGGGCCCGCTGGCAAGCCGGCTACTACACCCACTGCAGACAGGAGGCCGGGATGGTGG
 AGGGCCCTCCAGGACCTCAAGGTGGCCCGGCTACTACACCCACTGCAGACAGGAGGCCGGGATGGTGG
 AGTTCAGGAGGGCGTGGAGCTGGTGGACGTACGGGTGCCCTTTATCCGGCTGAGGATGACGACGAGAACAG
 CTGCTGGTGGAGGCCATCGACGTGCCCGCAGGCAGTCGCCACCCCTCGGCCGCCCTGGTAAACATCACCACATCAA
 GGAGCAAGCCAGAGACGTGGTCTTGTGAGCAGCCTGAGTTCTCGGTGAGCCGGGGACAGGTGGCCCGCATCC
 CTGTCATCCGGCGTGTCTGGACGCCAGTCCAGGTCTCTACCGCAGACAGGATGCCACCGCGCAGGGCAAC
 CGGGACTACATCCCCGTGGAGGGTGGACTCTCCCTGGGGGGCCAGGTCCAGGTCTGGCAGGAGCTCAGCAACCC
 GGAGCTGCAAGAAGTGTGACTCCCTCTGGGGGGCCAGGACCCACTCATCATCAGGGACCCAGATGAACCTGG
 AGTGGGGCCACCTGGGCCAGGACCCACTCATCATCAGGGACCCAGATGAACCTGGGAGGAGCTTCACG
 AGTCAGATGTTGTCATCACAGCCACCCCTCACGGCGACCTGGGCCGGCAGAACCCCAATGCTAAGGCCGCTGG
 GTCCAGGAAGATCCATTCAACTGGCTGCCCTTCTGGCAAGCCAAAGGGTACAGGGTAAAGTACTGGATT
 GTGACTCCGAATCCGAAGCCACCTGCTGACAGCAAGGTGCCCTCAGTGGAGCTACCAACCTGTACCC
 GACTATGAGATGAAGGTGTGCCCTACGGGGCTCAGGGCGAGGGACCTACAGCTCCCTGGTGTCTGCCGCACCC
 CCAGGAAGTGCCAGCGAGCCAGGGCGTCTGGCCTTCAATGTCGTCCTCCACGGTACCCAGCTGAGCTGG
 AGCCGGCTGAGACCAACGGTGAGATCACAGCCTACGAGGTCTGCTATGGCTGGTCAACGATGACAAACCG
 ACCTATTGGCCATGAAGAAAGTGTGGTGAACAACCTAAGAACCGATGCTGCTTATTGAGAACCTCGGGAG

FIGURE 11B

TCCCAGCCCTACCGCTACACGGTGAAGGGCGCAACGGGGCCGGCTGGGGGCCTGAGCAGGGAGGCCATCATCAACCT
 GGCCACCCAGCCCAAGAGGCCATGTCCATCCCCATCATCCTGACATCCCTATCGTGGACGCCAGAGCAGGGAGG
 ACTACGACAGCTCCTTATGTACAGCGATGACGTTCTACGCTCTCCATCGGGCAGCCAGAGGCCAGCGTCTCCGAT
 GACACTGAGCACCTGGTGAATGGCCGGATGGACTTTGCCTTCCCAGCAGCACCAACTCCCTGCACAGGATGACCAC
 GACCAGTGTGCTGCCTATGGCACCCACCTGAGGCCACACGTGCCACCGCGTCTGCTAACGCACATCCTCCACCCCTCA
 CACGGGACTACAACACTCACTGACCCGCTCAGAACACTCACACTCGACCACACTGCCAGGGACTACTCCACCCCTCACC
 TCCGTCTCCTCCACGACTCTCGCTGACTGCTGGTGTGCCGACACGCCACCCGCTGGTGTCTGCCCTGG
 GCCCACATCTCTCAGAGTGAGCTGGCAGGAGCCGGTGCAGCGCTGAGGGTACAGTGTGGAGTACCCAGC
 TGCTGAACGGCGGTGAGCTGCATCGGCTCAACATCCCCAACCTGCCAGACCTCGGTGGTGGAAAGACCTCTG
 CCCAACCAACTCCTACGTGTTCCGCGTGCAGGGCCAGAGCCAGGAAGGCTGGGCCAGAGCGTGGGTGTCACTCAC
 CATTAATCCCAGGTGACCCCGAGAGCCCACGTGTCCTGCCAGGCTCCGCTTCAACTTGAGCAGCTCCAGTG
 CCCCAGGCCGCTGGTGTTCAGTGCCTGAGCCCAGACTCGCTGAGCTGGAGCAGGCCACGGAGGCCAAT
 GGGGATATCGTCGGTACCTGGTGACCTGTGAGATGGCCAAGGAGGGCCAGCCACCCATTCCGGTGGATGG
 AGACAGCCCCGAGAGCCGGCTGACCGTGCCGGCTCAGCGAGAACGTTCAAGGTGCAGGCCAGGA
 CCACTGAGGGCTTCGGGCCAGAGCGCAGGGCATCATCACCATAAGGTCCCAGGATGGAGGCCCTCCGCACTG
 GGCAGCCGTGCCGGCTTCCAGCACCGCTGCAAAGCGAGTACAGCAGCATCACCACCACACCAGGCCAC
 CGAGCCCTCTAGTGGATGGCCGACCCCTGGGGCCAGCACCTGGAGGCAGGGCTCCCTCACCCGGCATGTGA
 CCCAGGAGTTGTGAGCCGGACACTGACCAACAGCGGAACCCCTAGCACCCACATGGACCAACAGTTCTCCAAACT
TGACCGCACCCCTGGCCACCCCGCCATGTCCACTAGGCGTCTCCGACTCCTCTCCGGAGCCTCTCAGCTAC
 TCCATCCTGCACTGGGGCCAGCCCACCCGATGCACAGAGCAGGGCTAGGTGTCTCTGGGAGGCATGAA
 GGGGCAAGGTCCGTCTGTGGCCAAACCTATTGTAACCAAAGAGCTGGGAGCAGCACAGGACCCAGCCT
 TGTTCTGCACTTAATAATGGTTTGCTACTG

FIGURE 12A

GCCTTCAACTACCATCCCACCCACCTGCTGAGGAGAAAAATTCTCAAGACTCAGAGCACACAGCCAGCACCAGAGGC
 CCC**ATG**ACCCCTGGACAGACCAGGGAGGGGGCCACCATGCTGAAGACATTCACTGTTGCTCTTCGATTCGGAT
 GAGTCTGGGTATGACATCGATAGTGTGACCCCTCAACCGGAGTTGGGATAGAGTCCAACCTACCCCCAGGCCCT
 GGGAGAACATCACGGTTGGTGCAGGCCAGCCCCCTCGGATATCAAGCAAGTCCCTGCTGCTGAAGGATAAGACACAA
 ATGACCTGGATCCGCCCTCCCACAAGACCTCCAAGTTCTTACAGGTTAGGGCCCTACTGAGTCCAATGCAGG
 TCTTACCGGTGCTGACTTGGAGGAGACAGGCTGGTCAAAGCCCAGTAAAGTTCTAGAGTTGGAGGCACCAGGCC
 AACTGCCAAGCCATCTCTGGATTCAAGGCTGAGACCCCCGCTCTTCTGGGTGTAATGTTAACATCCTCTGCCAT
 GGCTGGCTGCAGGATTGGTATTCATGCTGTTAAAGAGGGATATGCAGAGCCTGTTGATTACCAAGTCCAACCTGG
 GACAATGGCCATATTCTCCATTGACAACCTGACACCTGAGGATGAAGGGTTACATCTGCCCACTCATATCCAGA
 TGCTCCCCACCCCTGTGGTCAGAGCCCAGCAACCCCTGAAGCTGGTTAGCAGGACTCTACCCCAAACCAACTTTG
 ACAGCCCACATCTGGGCCCATCATGGCACCTGGAGAAAGCCTGAATCTCAGGTGCAAGGGCAATCTATGGAATGAC
 CTTTGCTCTAATGAGGGTTGAGAGACTTGGAGAAGTCTTACACAAAGAACATAAAAGAGGGCAAATTCT
 TCTTCCAGTCTTGAGAGATCCAAGATACTGGACATTACCTCTGTTTACTATGACGCATCATATAGAGGGTCACTC
 CTTAGTGATGCTCTGAAAATCTGGTAACTGACACTTCCCCAAGACCTGGCTACTGCTGGGCCAGTGTGCTGTGGT
 CCAAATGGGTCAAGATGTGAGCCTACGGTGTGAGGACAGTGGATGGTGGGCTTGCACTCTATAAGAAAGGAG
 AAGACAAACCACTTCATTTGGATGCCACCAGCATCGATGACAACACATCATCTTCTCAACAATGTAACCTAC
 AGTGTACTGGCATCTATAGTGTGCCACTATCTTCACCTGGAAGACCTCATTAGGATGCATCACACAACACTGT
 GGAGCTATGGTTGAGATAAGCCCCCAAACCCCTCCCTGTCAGCTGGCCAAGCACTGTGTTCAAGCTAGGAAAGG
 CCATCACCCCTCAGTGGCAGTATCTCATCCAGTACTGGAAATTCTCTGGAAATGGGAAGAAAGAGAACATTCCAA
 AGATTCTCAGTAAACGGAGACTTCATCATCAGTAATGTTGACGGGAAGGGCACAGGGACCTACAGTGCAGCTATCG
 CGTAGAGACACATCTAACATGTGGTCACATCGCAGTGGCCCTGAAGCTGATGGGCCAGCAGGCTATCTCACCT
 GGAATTACGTTCTGAATGAAGCTATCAGGTTGCTCTAATCATGCACTGGCTTGGCTGCTGGTAGTGCTGTGG
 ATAAGGTGGAAGTGTGGAGACTCAGAATCAGAGAACGCTGGTTGCTGGGAACAGCTCAAGGGGTACCATGCTCTT
 CATAGTCACGGCCCTCTGCTGTGGACTGTGCAATGGGTATTGATAGAAGAGACTGAAATAGTCATGCCAACCC
 CTAAGCTGAGCTGTGGCAGAGACCAACTTCTCTGGCCCGTGGAAAGAACTTAACCTCTGGTGCAGAACCC
 TCTGGCTCAACTAAGGAGTTGTGGCTGAAGGATGGGACCGGGTGATGCCACTGCCCGGCTCAGAGCAGGT
 CGGGCTGCCCTCCCCCTGGCGCCCTGACCCAGGCCACACCCGGGAGCTACCAACTGCCATTCATGGAGGAGATGG
 CTGATCGGAGGCCAGTGGAGGCACTTGAGCTGGTGGGACAGACATCTCCCCAAACCTGTCATTCTGCTCTCCCC
 ACAATCGGGGCCAGGAACACTCCGGTGCAGGATGGCTGCCAGGATGGGTTGCTGTATAAGGAGGG
 AGAGCAAGAACCTGTCAGCAACTTGGCTGTTGAAGAGAACGGCTTCTTACAATCCAGAGAACGGAGATAAG
 ACGAAGGCAATTACAGTGCCTGCACTCACACTGAAACACTCCCTCAAGTGGCTGAGCCAGTGGCTGGCAGG
 CTTGTCATAAAAGAAATGTACCCCTAAGCCCTCTCAAGACATGGCCAGCCCTGTTGTCACCCCTGGTGGCGAGT
 GACTTTCAATTGCTCCACCCCCCACCAGCATATGAGCTTATTCTTACAAGATGGAAGTGAATAGCATCCAGTG
 ACAGGTCTGGCAGTCCGGGGCCAGTGCAGCTCACTTCTAATCATTCGGTGGGCATTGGTGTGGAGGGAAAT
 TACAGCTGCCGATATTATGACTTTCTATCTGGCTGAGCCAGCGACCCCTGTGGAGCTGTTGACAGAATTCTA
 CCCCCAAACCACTCTCTGGCACAGCCAGGTCTGGTGGCTGGGAAGAGTGTGATCTGCCTGCCAACGGGA
 CTTTCAGGGCATGGGTTGCCCTCTGCAAGGAGGGAGCCATGTTCCCTACAGTTGGAGTGTCTCAGGGAAC
 TCAGCTGACTCCTCTCCACACTGTTGGAGCAGAGGACTCTGGAAACTATAGCTGTATCTACTATGAGACAAACAT
 GTCAAACAGGGGTCAATCTCAGTATGCCCTATGATCTGGGTACTGACACATCTCTAACGGCATGGTTGGT
 CTGAGCCAGTCTGTGGTCCATGGGCAGAATGTTACTCTCTGGTGCAGGGCCGGTCCATGGAGTAGGATAC
 ATTCTGCACAAAGAAGGAGAAGCCACTTCATGCAATGCAGCTGGGATCACCAGTAATGACGGGCATTCCCCATCAC
 CAATATATCTGGTACTAGCATGGGCGTTACAGCTGCTGCTACCACCTGACTGACCCAGTCTATCAAGATAAAC
 CTAGCAACACCCCTGAACTCTAGTCACAGGCTACTCCCCAAACCCAGCCTATTAGGCCAGCCTGGTCCATGGT
 GCCCTGGCGAAAATATGACTCTCAGTGTCAAGGGGAACGGCCAGACTCAACATTGTGCTGGTGAAGGAGGG
 TCAGGAGCCTTAGAGCAACAGAGGCAAGTGGGACTCAGGGCTGACTCTGGATGCCAGCAGTGGAGGGTGAAGACT
 CTGGGATCTATAGCTGTGTTATTATTGGACTCTACTCTGGGATCTGCAAGTCAAGTGGAGTGTGAGGGAGAC
 TGGGTGACTGATAAGCCCCCTAAACCCCTCTGTCAGGCCAGGCTGAGGTTAGGAAACTACAGCTGCCAGTACCC
 CCTTCAGTGCCTGGAGGACCCCTGCCAGGTGTTGAATTGTTCTAGAACATGATGGAGAAGAACCTCAGCAGTTT
 CAGAGGAGGGAGACTTGTGATCAACACGTAGAAGGAAAGGACTGGAAACTACAGCTGCCAGTACCC
 GCCTACCCCTGATATCTGGTCAAGGCCAGTGTGATCCCCCTGGAGCTGGTGGGGCAGCAGGGCTGTTGCTCAGGAGTG
 CACTGTAGGGAACATTGTCCGAAGTAGCCTAATCGTGGTGGTGTGAGCCTGGGGTAGTGCTAGCCATAGAGT
 GGAAGAAGTGGCCTCGACTGCCAACCAGAGGCTCAGAGACAGACGGAAAGAGACCAGACCATTGCCCTG

FIGURE 12B

AAGAGTGTAAACCAAGAAGGAGAACCGAGGCACCCCTGCCAATTCTCCTTCATCAACCTCTCAGAGAATCTCTGTGGAA
CTGCCCGTTCCAATATAATAATCTCCTCCTTACAAGAGCTTCCTCTCTCTCTTGCTCTCAGAGACCTATAAA
TCCAACCAGTTACCCCTGCAAGTCAGCCCCATCTGCTGTTCCCTGGTCTTAATCACCTGAGCTGGTAAAGGGGATT
CTGGGAGTTGAGAGCTCTGCCAGGGTGAGATGTTCCCTGAAGAGAGGTTCCCACCCCTGTAACTCCTCACTGTACT
GATTTACTGGCGCATGAAATTCTATTAAAAATGCATTCTGAATAAAAAGAGTATTCACTATTTAACTTCATTT

FIGURE 13A

CGGGAGCGGGGGAGCGGTGGCGGGCAGAGGGGGCGGCTCCAGCTCGGCTCCGGCTCGGGCTCGGGCTCCGGCT
 CGGGCTCCGGCTCCGGCTCCAGCTCGGGTGGCGGGAGCGGGACCGAGGTGGAGGCAGGGCGGAGAGGAGT
 GGGAGCAGCGGCCCTAGCGGCTTGCGGGGGACATCGGACCGACGGCCCTGGATAGGCAGAAGGAGTGGAGGCC
 TGGTGGCCGGCCCTTGGTGTGAGTATCCAGCAAGAGTGACCGGGGTGAAGAAGCAAAGACTCGGTTGATTGCTCG
 GGCTGTGGCTGGCTGTGGAGCTAGAGCCCTGGATAGGCCCTGAGCCAGCCCCAGGGAGGACGATGGTCCCCTGTG
 CCTGCACTGGTATGCTTGGTTGGCAGGCAGCCATGGTACAGCAAACCTGTCTTCATTAAAGCTCTGAGGA
 CCAGACTGGCTGTCAAGGAGGGTAGCCTCCTCGTGTGCCAAGCTACAGGAGAACCCAAGCCGCGCATCACATGGA
 TGAAGAAGGGAGAAAGACTAGCTCCAGCCTGAGGTATTGAGTTGATGATGGGGCAGGGTCAGTGTCTCG
 ATCCAGCCATTGCGGGTGCAGCGAGATGAAGCCATCTATGAGTGTACAGCTACTAACAGCCTGGGTGAGATCAACAC
 TAGTGCCAAGCTCTCAGTGTCAAGAGGAACAGCTGCCCTGGGTCCCTCATCGACATGGGCCTCAGCTGA
 AGGTGGTGGAGAAGGCAGCACGCCACCATGCTATGTGCCCGCAGGGGAAATCAGACCCCTGAGATTCTGGTT
 AAGGACTCCTCTGTAGACCCCTGCCACGAGAACCGCCGATCAAGCAGCTGCCCTAGGTGCTCAGGAGCTTGAGATAGA
 GAGCAGTGGAAATCGACCAAGGGCAAGTACGAGTGTGGCAGCAACTCGGCCAGGCACAGTACTCAGCCCC
 CGAACCTGTATGTGCGAGTGCAGCGCTGGCATGGTGCACCCATGCCCTACGTGAAGTGGATGATGGGGCCAGGAGC
 GGCAGCGTGAACCTGACATGCGTGGCAGTGGGTGCACCCATGCCCTACGTGAAGTGGATGATGGGGCCAGGAGC
 CACCAAGGAGGATGAGATGCAAGTTGGCGCAACGCTGGAGCTCAGCAATGTCGTACGCTCTGCCAAACTACAC
 GTGTGGCCATCTCTCGTGGCATGTCAGAGGCCACAGCCAGGTACAGTGAAGCTCTCAAAGCCTCCGATT
 GATCTTGTGGTACAGAGACAACGCCAGTGTACCCCTCACCTGGGACTCTGGGACTCTGGGACTCTGGAGCCTGTAACCTA
 CTATGGCATCCAGTACCGCGCAGCGGGCACGGAGGGCCCTTCAGGAGGTGGATGGTGTGGCCACCACCGCTACA
 GCATTGGCGGCCTCAGCCCTTCTCGGAATATGCCCTCCCGTGTGGCGTGAACAGCATCGGGCGAGGGCCGCC
 AGCAGGGCAGTGCAGGACGACGGAGAACAGGCCCTCAGGCCACCGCGCGTGCAGGACAGCATGTC
 CGCCAGCACCAGTGTGGTCACTGGGAGCCTCCGAGGAGCCAACGCCCTGGTGCAGGGGATACCGCGTACTATA
 CTCCGGACTCCCAGGCCCGAACGCCCTGGCACAAGCACACCGACGCCGGCTCCTCACGACCGTGGCAGC
 CTGCTGCCCTGGCATCACCTACAGCCTGCCGTGCTGCCCTCACGCCGTGGCGATGGCCCTCCAGGCCACCAT
 CCAGGTCAAGACGAGCAGGGAGTGCCTGCCAGCCAGGCCGACTTCCAGGCCAGGTGGAGTCGGACACCAGGATCC
 AGCTCTCGTGGCTGCTGCCCTCAGGAGCGGATCATCATGTATGAACCTGGTACTGGGCGCAGAGGAGCAAGAC
 CAACAGCACAAGGTCACTTCGACCCAACCTCCCTACACACTAGAGGACCTGAAGCCTGACACACTCTACCGCTT
 CCAGCTGGCTGCACGCTCGGATATGGGGTGGCGTCTTCACCCCAACATTGAGGCCGACAGGCCAGTCCACCC
 CCTCGGGCCCTCCAGAAGGTGATGTGTGAGCATGGGCTCCACCGGTGGGTAAGTTGGGTCGGCCGCT
 GCGCACAGCCGAAACGGCTTACACCCAGTACTCCGTGGGCCACAGGGCGGTGAGCAGGCCAGGAGACCGCGGGCGCA
 TGTGGTGGATGGCATCACCGTGGCAGCACCTCCAGTGGGCCCTGGGACCTGGTGGGCTGGAGAAGTGGACGGAGTACCGGGTGT
 GGGTGGGGCACACACAGACGTGGGCCCGGGCCAGAGCAGGCCGGTGTGGCGCACCAGTGGAGCAGTGGCC
 AGCGGGCTCCCGGAAGGTGGAGGTGGAGCCACTGAACCTCACTGCTGTGCTACTGGAGCTGCTGTCCC
 CAGCAAGCAGCATGCCAGATCCCGCGCTACCAAGGTACCTACGTGCCGTGGAGAATGGCGAGCCCCGTGGACTCC
 CCATCATCCAAGACGTATGCTAGCCAGGCCAGTGGCGCCAGAGGAGTCCGAGGACTATGAAACCACTATCAGC
 GGCCTGACCCCGGAGACCACCTACTCCGTTACTGTTGCTGCCTATACCAACCAAGGGGATGGTGCAGGCCAGCAAGCC
 CAAAATTGTCACTACAACAGGTGCACTCCAGGCCGCCACCATGATGATCAGCACACGCCATGAACACTGCGC
 TGCTCCAGTGGCACCCACCAAGGAACCTGGCTGGCAGCTGCTGGCTACCGGCTGCACTACTGCCGGCGACGGAG
 GCGCGGCCAACACCATAGATTGGCAAGGATGACCAGCACTTCACAGTCACCGGCTGCACAAGGGACCACCTA
 CATCTCCGGCTTGCTGCCAAGAACCGGGCTGGCTGGGTGAGGAGTTCGAGAAGGAGATCAGGACCCCGAGGACC
 TGCCCAGCGGCTCCCGAAACCTGCAATGTGACAGGACTGACCACGTCTACAGAACTGGCTGGGACCCGCA
 GTGCTGGCGGAGAGGAACCGCGCAGTACAGCTACACCGTGGTGTGGAGACATCAACAGCAACAGGAGCTGCA
 GAACATCACGACAGACACCCGCTTACCCCTACTGGCTCAAGCCAGACACCACTACGACATCAAGGTCCGCGCAT
 GGACCGAAAGGCTGGCCACTCAGGCCAGCATCCAGTCCCGGAGCATGCCGGTGGAGCAAGTGGTGTGCAAG
 AACTTCCGGTGGCGCTGCAATGAAGACGTCTGTGCTGCTCAGCTGGGAGGTGGCGACTCTATAAGTCAGCTG
 GCCCTTAAGATTGTACAATGGGAGAGTGTGGAGGTGGACGGGCACTCGATGCCAGGAAAGCTGATGCCAGACCTGC
 AGCCCACACAGAGTACTGTTGTGCTGATGAACCGTGGCAGCAGCCAGGGGCGTGCAGCACCTGGTGTCCATC
 CGCACAGCCCCGACCTCTGCCCTCACAGCCGCTGCCCTGCTACATAGAGGAGCCGCTTCGATCTCTC
 CATGCCCTATGTGCAAGACCCCTCGCTGTGCAAGGTGGAGCACACCGAGGAACGGAGCTGGAGGAGCTCTAGAAG
 GCATGCTGACGCCAAGGTGGAGCACACCGAGGAACGGAGCTGGAGGAGCTCTAGAAGCCATGAGCAAGGCCGA
 GAGGAGCAGCGGGCGGGCGGGCAGGAGAACGTCTGAAGCCATATGTGGCTGCTCAACTGGATGTGCTCCGGACT
 GACCTTACCTGGGGACAAGAAACTACCGGGCTTCTACAACCGGCCCTGTCTCCGGACTTGAGCT

FIGURE 13B

ACCAAGTGTCTTGCTTGCCTCCTGAAGGAACCCATGGACCAGAAGCGCTATGCCCTCAGCCCCACTCGGATGAG
ATCGTGGTCCAGGTGACACCAGCCAGCAGCAGGAGGAGCCGGAGATGCTGTGGGTGACGGGTCCTGCTGGCAGT
CATCCTCATCATCCTCATTGTCATGCCATCCTCTTGTCAAAAGGAAAAGGACCACTCTCGTCCCTAAGGATG
AGCAGTCGATCGGACTGAAGGACTCCTGCTGGCCACTCCTGACCCCTGTGGAGATGCAGGCTCAACTACCAAG
ACCCCAAGGTATCGGAGACCACCCACCCATCCCCATCACCAGACCTGGGGACAACATCGAGGCCCTCAAAGCCAACGA
TGGCCTCAAGTCTCCCAGGAGTATGAGTCATCGACCTGGACAGCAGTCACGTGGGAGAATTCAAACCTGGAGG
TGAACAAGCCCAAGAACCGCTATGCGAATGTCATCGCCTACGACCACTCTGAGTCATCCTACCTATCGATGGC
GTCCCCGGGAGTGAATCATCAATGCCAACTACATCGATGGCTACCGCAAGCAGAATGCCATCGCCACGCAGGG
CCCCCTGGCCAGGACCATGGGCAATTCTGGAGAATGGTGTGGGACAGCGCACGCCACTGTGGTCATGATGACAC
GGCTGGAGGAGAAGTCCCAGGAAATGTGATCAGTACTGGCCAGCCGGAGACACTGTGGCTTATTCTAG
GTGACCCCTGTGGACACAGTGGAGCTGGCCACATACACTGTGCGCACCTTCGACTCCACAAAGAGTGGCTCAGTGA
GAAGCGTGAAGCTGGCTCAGTTCAAGTTCATGGCCTGGCCAGACCATGGAGTTCTGAGTACCCAATCCTCGG
CCTTCCTACGACGGGCTCAAGGCTGCAACCCCTAGACGCAAGGGCCATGGTGGTCAGTGCAGGCCGGCTGGGC
CGCACCGGCTGCTCATCGTATTGATGCCATGTTGGAGCGGATGAAGCACGAGAAGACGGTGGACATCTATGGCC
CGTGAACCTGCATGCGATCACAGAGGAACATGGTGCAGACGGAGGACCACTACGTGTTCATCCATGAGGCGCTGC
TGGAGGCTGCCACGTGCGGCCACACAGAGGTGCCCTGGCCGAAACCTGTATGCCACATCCAGAAGCTGGCCAAGTG
CCTCCAGGGGAGAGTGTGACGCCATGGAGCTGACTCAAGTTGCTGGCAGCTCCAAGGCCACACGTCCCCTT
CATCAGGCCAACCTGCCCTGCAACAAGTCAAGAACCGGCTGGTAACATCATGCCCTACGAATTGACCCGTGTT
GTCTGCAGCCCATTGGTGTGGAGGGCTGTGACTACATCAATGCCAGTCTCTGGATGGTTAGACAGCAGAAC
GCCTACATAGCTACACAGGGCCTCTGGCAGAGAGCACCAGGGACTTCTGGCGCATGCTATGGAGCACAATTCCAC
CATCATCGTATGCTGACCAAGCTCCTGGAGATGGCAGGGAGAAATGCCACCACTGGCCAGCAGAGCGCTCG
CTCGCTACCAAGTACTTGTGACCCATGGCTGAGTACAACATGCCCACTGATATCTGCTGAGTCAAGGTC
ACGGATGCCCGGGATGGCAGTCAGGACAATCCGGCAGTCCAGTTCACAGACTGGCCAGAGCAGGGCGTGCCAA
GACAGGGAGGGATTCTGACTTCATGGGCAGGTGCTAAAGACCAAGGAGCAGTTGGACAGGATGGCCTATCA
CGGTGCACTGCTGGCTGGCGACCGGGGTGTTCATCACTCTGAGCATGCTCTGGAGCGCATGCGCTAT
GAGGGCGTGGTGCACATGTTCAAGACCGTGAAGAACCTGCTACACAGCGTCTGCCATGGTGCAGACAGAGGACCA
GTATCAGCTGTGCTACCGTGGCGCCCTGGAGTACCTGGCAGCTTGGACACTATGCAACGTAACGTTAACTACCGCTCCCC
CTCCTCCGCCACCCCGCCGTTGGGCTCCGGAGGGGACCCAGCTCTGGTACAGCCATACCGACATCGTCCAGGCCCTC
CTACGCAGATGCTGCACTGGCAGAGCACAGCCCACGGGGATCACAGCGTTCAAGGAACGTTGCCACACCAATCAGA
GAGCCTAGAACATCCCTGGCAAGTGGATGGCCAGCAGGAGGACTGTGGCCCTCTGTTCTGACACAGGCCACCTG
GAGCCGCTTCAAGCTCTGCGCTCCGCAATTCTCATGCTTCTCATGGGTGGGTTGGGCAATCCGTA
CCTTTTAATACATTAAGTGGGTTAGACTGAGGGATTTCAGGCTTCTCCCTGTGATTTTCTTCTGGAATCCGTA
TCTGAGAATGGGCACGTAGGGGTTGGGTTATTTGTTGTTTTTTTTGTTGATGACTCTGCTGA
AGGACAGAACATTGCCCTCCGTGCAAGAGCTGGGCTGCCAGCCTGAGCGAGGCTCGGCCGTGGCCGGAGGCA
GTGCTGATCCGGCTGCTCCCTCAGCCCTCAGACGAGATCCTGTTCAAGCTAAATGCAGGGAAACTCAATGTTTT
TAAGTTTGTTCCTTAAAGCTTTTTAGGCCACATTGACAGTGGTGGGGGGAGAAGATAGGGAACACTC
ATCCCTGGTGTCTATCCCAGTGTGTTAACATTACAGCCAGAACACAGATGTCCTGGGAGAGCCTGGCAA
GGCATTCCATACCACGTGTTGCAAAAGGTTAAAACAAAAACAAAAACACAAAAAATAAAAACAAAAAAAC
AAAAAAACCAAAAAAAAAAAAGAGTCAGCCCTGGCTCTGCTCAAACCCCTCAAGAGGGAAAGCAACTCCG
TGTGCCTGGGTTCCCGAGGGAGCTGCTGGCTGACCTGGGCCACAGAGCCTGGCTTGGTCCCCAGCATGCAAGTA
TGGTGTGGTTGTTGAGGCTGTTGGGCTGGCTGTTGGCAAGGTGAATAGCACAGGTTAGGGTGTGTCACACC
CCATGCACCTCAGGGCCAAGGGGGCGTGGCTGCCCTTCAGGTCCAGGCGAGTGGCCTGGTAGCACATGCTGT
CCTCAGAGCAGGGGCCAGATGATTCTCCCTGGTTGCGCTGTTCAAAAGCCCCGATAATCGCTTCTTCCA
CTCCAAGATGCCCTCATAAACCATGGCAAGACTACTGGACTTCTATCAATGGTACTCTAATCAGTCTTATTAT
CCCAGCTGCTGAGGGCGAGGGAGAGCGCCTCTGGCAGCGCTATCTAGATAGGTAAGTGGGGCGGGGAA
GGGTGCACTGCTGTTAGCTGAGGGACGTGGTGCCGACGTCCCCAAACCTAGCTAGGTAAGTCAAGATCAACATT
CCAGGGTTGGTAATGTTGGATGATGAAACATTCAATTACCTTGATGCTAGTGTAGAGTTCAACTGTTGTA
CACAGTCGTTCTATTGTTAAGAAAAACTACAGCATCATTGCAATTCTGATGGTAATAAATTGAATAATC
AGATTCT

FIGURE 14

GGAGAGGGTGCAGGGCGAATCCGAGCCGAGCGAGAGGAATCCGGCAGTAGAGAGCGGACTCCAGCCGGCGGACCCCTGC
 AGCCCTCGCTGGACAGCGCGCTGGCAGGCAGCCAAAGAGAGCATCGAGCAGCGGAACCGCGAAGCCGGCC
 GCAGCCGCAGCCCGCAGCCTGCCGCTCTCCGCCGCGGTCCGGCAGCATGAGGCGCGCGCTCTGGCTCTG
 GCTGTGCGCGCTGGCGCTGAGCCTGCAGCTGCCCTGCCGCAAATTGTGGCTACTAATTGCCCTGAAGATCAAG
 ATGGCTCTGGGATGACTCTGACAACCTCTCCGGCTCAGGTGCAGGTGCTTGCAGAAGATATCACCTGTCACAGCAG
 ACCCCCTCCACTTGAAGGACACGCAGCTCCTGACGGCTATTCCCACGTCTCCAGAACCCACCGCCTGGAGGCTAC
 AGCTGCCTCCACCTCACCCCTGCCGGCTGGAGAGGGGCCAAGGAGGGAGAGGCTGTAGTCTGCCAGAAGTGGAGC
 CTGGCCTCACCGCCGGAGCAGGAGGCCACCCCCCGACCCAGGGAGACCACACAGCTCCGACCACATCAGGCC
 TCAACGACCACAGCCACCACGGCCCAGGAGCCGCCACCTCCCACCCCCACAGGGACATGCAGCCTGCCACCATGA
 GACCTCAACCCCTGCAGGACCCAGCCAAGCTGACCTTACACACTCCCCACACAGAGGATGGAGGTCTCTGCCACCG
 AGAGGGCTGCTGAGGATGGAGCCTCCAGTCAGCTCCAGCAGCAGAGGGCTCTGGGAGCAGGACTTCACCTTGAA
 ACCTCGGGGAGAATACGGCTGTAGTGGCGTGGAGCCTGACCGCCGAACCAGCTCCAGTGGATCAGGGGCCAC
 GGGGGCCTCACAGGGCTCTGGACAGGAAAGAGGTGCTGGAGGGTCATTGCCGGAGGCCTCGTGGGCTCATCT
 TTGCTGTGCTGGTGGTTCATGCTGTACCGCATGAAGAAGAAGGACGAAGGCAGCTACTCCTGGAGGAGCCG
 AAACAAGCCAACGGGGGCCCTACCAGAACAGCCACAAACAGGAGGAATTCTATGCCTGACCGGGAGCCATGCC
 CCCTCCGCCCTGCCACTCACTAGGCCCCACTTGCTCTTGAAGAACTGCAGGCCCTGCCCTCCCTGCCACC
 AGGCCACCTCCCCAGCATTCCAGCCCCCTGGCGCTCTGCCACGGAGTCGTGGTGTGGAGCTCCACTCT
 GCTTCTCTGACTTCTGCCCTGGAGACTTAGGGCACAGGGTTCTCGCATAGGACCTTCCACCACAGCCAGCACCT
 GGCATCGCACCATTCTGACTCGGTTCTCAAACCTGAAGCAGCCTCTCCCAGGTCCAGCTCTGGAGGGAGGGGA
 TCCGACTGCTTGACCTAAATGGCCTCATGTGGCTGGAAGATCTGGGGTGGGGCTGGGCTCACACACCTGTAG
 CACTTACTGGTAGGACCAAGCATCTGGGGGTGGCGCTGAGTGGCAGGGACAGGAGTCACTTGTTCTGGGG
 AGGTCTAATCTAGATATGACTGTTTGACATGTTCTCATGTTCTGTACAGCCCAGTAGACCTTGTAA
 CTTCTGAGGTAAGTAAAGTAAAGTGTGATTCGGTATCCCCCATCTGCTTCCCTAATCTATGGCGGGAGACAGC
 ATCAGGGTTAAGAAGACTTTTTTTTTAAACTAGGAGAACCAAATCTGGAGCCAAATGTAGGCTTAGTTG
 TGTGTTGCTCTTGAGTTGTCGCTCATGTGTGCAACAGGGTATGGACTATCTGCTGGTGGCCCGTTCTGGTGG
 TCTGTTGGCAGGCTGGCCAGTCCAGGCTGCCGTGGAGAACGCCGTGTGAGAAGTGAATGCTGGACTCAGCCT
 TCAGGGCCATGCTGAGGCCTGGCCGTGCCACGTTGGAGAACGCCGTGTGAGAAGTGAATGCTGGACTCAGCCT
 CAGACAGAGAGGACTGTAGGGAGGGCGCAGGGGCCTGGAGATCCTCTGCAGACCACXCCGTCTGCCGTGCGC
 CGTCTCAGGGCTGCTCCCTGGAAATTGACGAGGGTGTCTGGCAGAGCTGGCTTGAGCGCCTCCATCCA
 AGGCCAGGTTCTCCGTTAGCTCCTGTGGCCCCACCCCTGGGCCCTGGGCTGGAATCAGGAATATTTCCAAAGAGTGA
 TAGTCTTTGCTTTGGCAAAACTCTACTTAATCCAATGGTTTCCCTGTACAGTAGATTCCAAATGTAATAA
 ACTTTAATATAAAGT

FIGURE 15

GCTCCGGCCAGCCGGTCCAGAGCGCGCAGGGTCTGGGGAGCTCCGCCAGGCTGCTGGTACCTGCGTCCGCCCGGC
GAGCAGGACAGGCTGCTTGTTGTGACCTCCAGGCAGGCAGGCCATCCTCTCCAGAATGAAGATCTTCTGCCAG
TGCTGCTGGCTGCCCTCTGGGTGTGGAGCGAGCCAGCTCGCTGATGTGCTTCTCTGCTTGAACCAGAAGAGCAAT
CTGTACTGCCTGAAGCCGACCATCTGCTCCGACCAGGACAACACTGCGTGACTGTGCTGCTAGTGCCGGCATTGG
GAATCTCGTGACATTGGCACAGCCTGAGCAAGACCTGTCCTGGCCCTGCCCATCCAGAAGGCGTCAATGTTG
GTGTGGCTTCATGGCATCAGCTGCTGCCAGAGCTTCTGTGCAATTTCAGTGCAGGCCATGGCGGGCTGCCGGCA
AGCGTCACCCCTGCTGGGTGCCGGCTGCTGCTGAGCCTGCTGCCGCCCTGCTGCCGGTTGGCCCCTGACCGCCAG
ACCCCTGCCCCGATCCCCCAGCTCAGGAAGGAAAGGCCAGGCCCTTCTGGATCCCACAGTGTATGGGAGGCCCTGA
CTCCTCACGTGCCCTGATCTGTGCCCTTGGTCCCAGGTCAAGGCCACCCCTGCACCTCACCTGCCCAAGCCCCCTGC
CTCTGCCCAAGTGGGCCAGCTGCCCTACTTCTGGGTGGATGATGTGACCTCCTGGGGACTGCCGAAGGGACG
AGGGTTCCCTGGAGTCTACGGTCCAACATCAGACCAAGTCCATGGACATGCTGACAGGGTCCCCAGGGAGACCGT
GTCAGTAGGGATGTGCTGGCTGTGACGTGGGTGTGCAGTGCACGTGAGAGCACGTGGCGCTCTGGGGCCA
TGTTTGGGGAGGGAGGTGTGCCAGCAGCCTGGAGAGCCTCAGTCCCTGTAGCCCCCTGCCCTGGCACAGCTGCATGC
ACTTCAAGGGCAGCCTTGGGGTTGGCTAGGCCACTCCGGTCTAGGCCAAATCCAGCCAGTCCTG
CCCCAGCCCACCCCCACATTGGAGCCCTCTGCTGCTTGGCTCAAATAATACAGATGTCCCC

FIGURE 16

CAGCAGGTACAGCCCCCTCGAGGCAGACAGCGGGCCCCGCCGCACCAGAGCAGTGGTACAGGCATGGATGGGAAGAAAT
 GCAGCGTATGGATGTTCTACCTCTTGTATTACTTGTAACTTCAGCTGGATTGTGGATAGTATACTTCATAGCT
 GTGGAAGATGACAAAATTTCAGCTGAAAGGAAACCTGGTGTGAAGCATGCACCATAAAGCAT
 TGCAGGTGATGATCCTCCTGCAAGCTGTGTTAGTCAAGTTATGAACATGGCAGCCTCCTAGCCCTGTGGTAG
 CTGTTCTCGCCTCATACAACTGAAACCGAAGGTTAAACCCGTGGCTGAATATTAGTGGATTGGTGGCTGTGTT
 CTGGCTCCTCGGAATGACCTTACTTGGTAATTTCAGCTCACAAATGATGAAGAAATCCATAACGTGGAACTTC
 CTTGACCTTGGATTGGCACATTGACCTGCTGGATCCAGGCTGCCTGACACTCAAGGTCAACATCAAGAATGAAG
 GACGGAGAGTTGGAATTCCACGGGTTATTCTGTCGGCATCTATCACTCTGTGTTCTACTTCATCCTCATG
 GCCCAAAGCATCCACATGTATGCAGCCAGGGTCCAGTGGGCCTGGTCATGTGCTTCTGTCTTATTGGCACCTT
 TGCCGTGGAGTTCCGGCATTACCGCTATGAGATTGTTGCTCTGAGTACCAAGGAGAATTCTCTAAGCTCAGAAA
 GCCTGTCAGAAGCTCTGAATATCAGACTGACCAGGTGTAAACCATCAGTTTCTTGCTGGTGGAGGTGGGTGTA
 CAGTGGGGAGGGCCAGTAGGACACACTCACAGGACTTGACATAGAACCTCATTCACACACACACACACAC
 ATTCAATGCCACATTGCCAAATGAGCTTTCAGGGCGAGTTATTCTTTAATGAAAAAGCACAAGCCCTATGTG
 CGAAATACACGCTGTTACACTGAAAATATATGCACGACAGAGCAAGAAGCTGTGCATGATCACTTCTTATCCGCC
 CCTTCCCAGCACTCCCTCTTCCATTCTCCACATGTCATCAAGCACCCCTACCGAGTAGGGCAGGCCAAATGTT
 CCTTGGAGTAATGCCAACTCCGACGTTGCCTTCAGGTCAAAGGGCTGGAACCAGCTCGTGAGGAAGTTCTGAA
 TCTGGCACTAATATTCTTGAGTGGATAATAGTGTATCATAGAATAGGACGGAAATTGTATTGAGATGTGACCTGTG
 TCGCCTGTGGAAAGGCATAGTGAGAAGAACCTTCCCACGAAAGCCCCCTCATCGTTGTTAGTGGTGGCTGTG
 GATCCCAGGAGAGACATATGCCACAGACTGTGAGAGCAAAGCCGCCGTGTGATCTGGACTTGATGCACTGTGACT
 GAGAATGATTCCAAATGTGAATATGTGTAGGGACGTGGCTATCAGGCCTGGAAACAAGATGGGGCAGTGAAGGTA
 TGGTTAGTGTGTTGCTTCATAGTATGCCATGTACAATGTTTATATTCTAGTTCTTTAAGTAACCTACCATGA
 GTCTCTAAGCCTCATGGACAAAGATGTAGACCAAATGCAAGAGCTGAGCTTGTGCTTGGGTTCAACCATGATCAA
 GAAAAACTGAGGTCACCTGCAGGCTTACGTGGAGCTAACAGACAATATC

FIGURE 17

CTGCCTCCACTGCTCTGTGCTGGGATCATGGAACTTGCACTGCTGTGGGCTGGTGGTGTGGCTGGTGTGATTCC
 AATCCAGGGCGGGATCCTGAACCTGAACAAGATGGCAAGCAAGTGACTGGGAAATGCCATCCTCTCCTACTGGC
 CCTACGGCTGTCACTGCGGACTAGGTGGCAGAGGCCAACCCAAAGATGCCACGGACTGGTGTGCCAGACCCATGAC
 TGCTGCTATGACCACCTGAAGACCCAGGGTGCGGCATCTACAGACTATTACAGATACAACTTTCCAGGGAA
 CATCCACTGCTCTGACAAGGAAGCTGGTGTGAGCAGCAGCTGTGCCTGTGACAAGGAGGTGGCCTCTGCCTGA
 AGCGAACCTGGACACCTACCAGAAGCGACTGCGTTCTACTGGCGCCCCACTGCCGGGGCAGACCCCTGGTGC
TAGAAGCCCACACCCCTACCCGTTCCTCAGCATGGAGCCTGGCATCCCCACCTCAGTATCTAACCTGAACCAGC
 CTGGCTTTCAAACACTCCGGGGGAGGTAGTCCCAGCCTCCCCGGAACCCTTACCAATGCCTCTGACCTTCTG
 AAGCTTCCGAATCCTCCCAGTTGAGGCAGTAGCTGTGCTGAGGGTGGATGGAAATCTTGGGAGAAGCCAAAG
 CAAGGGAGCCCTCAGAGGTGGTTTGGACCAAAAGCATCGGGTGGGGGAGGGTCTGCCGTGCCCCACCTGCT
 GGCCCCCTGTCCCTCACCCCTCCAATATAGTCTCGAGCTACAACTGCAGCAGCCACTATAAAGGCAATAT
 TGATCTTCTGCCATGTGGCTTATCTTAAACCTCAAGGCCCTCCACTGTCCTAAGATAAGCCTCTCATAGG
 CACTGGGGACCCTCGCACAGTCTGCCATGTGACCCCTCCCCAGGCAAAGCTCTGAAGTCCCTGCAGGTGGAGGCCAT
 GCCTGTCTAAACTCAGTTGCATCCCTGGTGCCCAAAGCAACACCAGAACGAGCACTCCATAAATCCTTCTT
 GGGTGAAGCCTAGACAAGGCCCAGGTCTTGGGCTCCAGGCACCAGAGCCTGAGTACTTCTCCGTCCAGG
 CATTGGCTCAGGGTAGAATTACAAGGGCTACTGAATGGCTATTACTTCATCAGACTGATCCCCACCTCTCAGG
 TCAAAGGGCTACTTTCTGGAAGTCTCCCAGGCTGACTCCCTCCGTACTGCAAGGGCTCACTCCCTCCCAAG
 CTCCCCAATGCTCATGGCTTGCCGCTTACCTAGGTGGCAATGGAACTTCTGTATCTCCCC
 AACTAGACTGGAGCCCCGAAGGATGGAGACCATGTCTGTGCCCATCTGTTTCCCCTGTTTCCCACATACTAGGT
 GCTCAATTCTGCCTGTGAATGGCTGAGCCATAATGGATACACAGAGGTTGCAGCAGATGGGTGGGTACCTCAC
 CCAGATATCTCCAGGCCAAGGCCCCTCCCTGAGGTGAGGCCAGGGTTGGCAGCCAACTGCTCCAACTGCC
 CTTCCCCTAAATACTGCCCTGGCTAGGTGGGAGGTGCCTTCCCCCTGCCCCACCTCTCCACCAAGAGGCCACGT
 CACTCATGGCCAGGAGAGTGAACACATGGAGGGGTACAATTGCCAGCTCCCCGTGTCTGTGCAGCATTGTCTGGGTT
 GAATGACACTCTCAAAATTGTCCTGGATCGGCTGAGGCCAGGCTCTCCGTGAACCACCTCTGCTGGCTGA
 CCCCTGGCTATCCAGTTTCTGGTTCCCTCAAGGGTTCTCCAGAAAGTACCCCTCAGTAAAGCATTTGCACA
 AGAAAAAAAAAAAA

FIGURE 18A

GGCTGAAAGAGCCTGAGCTGTGCCCTCCATTCCACTGCTGTGGCAGGGTCAGAAATCTGGATAGAGAAAACCTT
 TGCAACGGGAATGTATCTTGTAAATCCCTAGCACGAAAGACTCTAACAGGTGTGCTGTGCCAGTCACCAACCA
 GCATATCCCCCTCTCCAAGTGCACACCCCAGCAAAATGAAGAGGAAGCAACACAGGTGGAGACTCAGCCTGAGA
 AATGGTCTGTCAGCACACCCAGAGCTACCCAAACAGATTCCATGGAGTTCTGAATTCCAGGGTGGCGGATAT
 TCCAATAAAGCCATGTATATCCGTGTATCCTATGACACCAAGCCAGACTCACTGCTCCATCTCATGGTAAAGATTG
 GCAGCTGGAACCTCCCAAGCTTAAATATCTGTGCATGGAGGCCTCCAGAACTTTGAGATGCAGCCAAAGCTGAAAC
 AAGTCTTGGAAAGGCCTGATCAAGGCTGCTATGACCACCGGGCTGGATCTCACCGGGGTGTCAGCACAGGT
 GTTATCAGCCACGTAGGGATGCCCTGAAAGACCACTCCTCAAGTCAGAGGCCGGTTGTGCTATAGGAATTGC
 TCCATGGGCATCGTGGAGAATAAGGAAGACCTGGTGGAAAGGATGTAACAAGAGTGTACAGACCATGTCCAACC
 CTCTAAGTAAGCTCTGTGCTCAACAACTCCACACCCACTCATCTGGCTGACAATGGCACCCCTGGCAAGTAT
 GGCGCCGAGGTGAAGCTGCGAAGGCTGTTGGAAAGCACATCTCCCTCCAGAAAGATCAACACAAGACTGGGGCAGGG
 CGTGCCCTCGTGGGTCTGTGGAGGGGGCCCTAACGTGGTGTCCATGCTCTTGGAAATACCTGCAAGAAGAGC
 CTCCCCATCCCTGTGGTGTGATTGTGATGGCAGCGACGTGCCGACATCTGTCTTGGCACAAGTACTGTGAA
 GAAGGCGGAATAATAATGAGTCCCTCAGGGAGCAGCTTAGTACCATCAGAAACATCTGCTTGGCACAAGTACTGTGAA
 ACAATCACATCAGCTGTTGCAATTATAATGGAGTGCATGAAGAAGAAAGAAACTCGTCACTGTGTTGAGATGGG
 CTGAGGGCCAGCAGGACATCGAGATGCAATTAACTGCCCTGCTGAAAGGAACAAACGTATCTGCTCCAGATCAG
 CTGAGCTTGGCAGCGCTGGAAACCGCGTGGACATAGCACCGAAGCCAGATCTTGTCTTGGGCCCCACTGGACGCC
 CCTGGGAAGCCTGGCACCCCCGACGGACAGCAAGCCACGGAGAAGGAGAAGAACCCATGCCACCAACCAAGG
 GAGGAAGGAGAAAGGGAAAGGCAAGAAGAAAGGGAAAGTGAAGAGAGAAGTGGAGGAAGAAACTGACCCCCGGAAG
 ATAGAGCTGCTGAACCTGGGTGAATGCTTGGAGCAAGCGATGCTAGATGCTTGTAGATGCTTAGATCGTGTGACTTGT
 GAAGCTCTGATTGAAAACGGAGTGAACATGCAACACTTTCTGACCATTCCGAGGCTGGAGGAGCTATAACACAA
 GACTGGGTCCACCAAACACACTCATCTGCTGGTGAGGGATGTGAAAAAGAGCAACCTCCGCTGATTACACATC
 AGCCTCATAGACATCGGGCTCGTGCCTGGAGTACCTCATGGGAGGAGCCTACCGCTGCAACTACACTCGAAAAACTT
 TCGGACCCCTTACAACAACCTGTTGGACCAAAGAGGCTAAAGCTCTAAACTCTGGGAATGGAAGATGATGAGC
 CTCCAGCTAAAGGGAAAGAAAAAAAAAGAAAAGGAGGAAGAGATCGACATTGATGTGGACGACCTGCCGTG
 AGTCGGTTCCAGTATCCCTCCACGAGCTGATGGTGTGGCAGTGTGATGAAACGCCAGAAAATGGCAGTGTCC
 CTGGCAGCGAGGGGAAGAGAGCATGGCCAAGGCCCTGGTGGCCTGCAAGCTCTACAAGGCCATGCCACGAGTC
 CCTGGAGACTGATCTGGTGGATGACATCTCCAGGACTGGATAACAATTCCAAGACTCGGCCAGCTTGCTTGGAG
 TTATTAGACCAGTCTATAAGCATGACGAGCAGATCGCTATGAAACTCTGACCTACGAGCTGAAAAGTGGAGCAA
 CTCGACCTGCCCTAAACTGGCGTGGCAGCAGGAAACACCGGGACTCTATTGTCACACCTGCAAGCCAGATGCTG
 CCGATATGTGGATGGGAAGACTGCGGATGCGGAAGAACCCGGCCTGAAGGTTATCATGGGATTCTTCTACCCCC
 ACCATCTTGTGGAATTTCGCACATATGATGATTCTCGTATCAAACATCCAAGGAAACGAGGATGGCAAAGA
 AAAAGAAGAGGAAATACGGATGCAATGCAGATGCTGGCTCAAGAAAGGGGATGAGGAGAACGAGCATAAAAAC
 AGAGAAGTATTCCATCGAACAAAGATCTGTGAATTCTATAACGCCCTATTGTCAGTTCTGGTTTACACAATA
 TCATACTGGGCTACCTGCTGTTAACTACGTCTGGTGCGGATGGATGGCTGGCCCTCCAGGAGTG
 GATCGTCATCTCTACATCGTGAGCCTGGCGTTAGAGAAGAGATACGAGAGATCCTCATGTCAGAACCCAGGCAA
 ACTCA
 GCCAGAAAATCAAAGTTGGCTCAGGAGTACTGGAACATCACAGATCTGTGGCCATTCCACATTGATTGGA
 GCAATTCTCGCTACAGAACCCAGCCATGGGCTATGGCGGGTGTACTGTGTGGATATCATCTCTGGTA
 CATCGTCTGGACATCTTGGTGTCAACAAGTATCTGGGCCATACGTGATGATGATTGGAAAGATGATGATCG
 ACATGCTGTACTTGTGGTGTACATGCTGGTGTGCTCATGAGTTCCGAGTAGCCCGTCAAGCCATTCTGCATCCA
 GAGGAGAAGCCCTTGGAAACTGGCCGAAACATCTTCTACATGCCCTACTGGATGATCTATGGAGAGGTGTTGC
 AGACCAAGATAGACCTCTACGCCATGAAATTAACTCTCTTGTGGTGTGAGAACCTATATGATGAGGAGGGCAAGCC
 TTCTCCCTGTATCCCCGGCCCTGGCTCACTCCAGCACTCATGGCGTGTATCTACTGGTCGCCAACATCTCTG
 GTGAAACCTGCTGATGTTGCTCAACAAACTTCTTGTGAAACTAAATCAATATCCAACCCAGGTGTTGGAAGTCCA
 GCGATATGAGCTGATGACATTTCATGACAGGCCAGTCCCTGCCACCGATGATCATTTAAGCCACATCTACA
 TCATCATTATGCGTCTCAGCGGCCGCTGAGGAAAGAGAGAACGGGACCAAGAGGAACGGGATCGTGGATTGAAG
 CTCTTCTTAGCGACGAGGAGCTAAAGAGGCTGATGAGTGTGAGGAGCAGTCAGGAGCACTTCCGGGAGAA
 GGAGGATGAGCAGTCAGCCAGCGACGCGATCCGGTCACTCTGAAAGAGTGTGAAAGATGATGAGG
 TGGAGAAATCAATGAAAGAGAAACTTTATGAAAACCTCCCTGCAAGACTGTTGACCTTCGACTGCTAGCTAGAA
 GAATTATCTAACAGAACGGTGAATGCTTGTGAAAGGAAACTCTTGTGAGGAGAACGAGCAGCATCAAGCGT
 GATGGCTACAGCTGCTGAGGAGAACGAGTATCTCTCCGGCAAAGCAGCATCAAGCGTGTGAGGAGCAG
 ATCGATATCATTAAACGGAGAAGAGTTATTGAGGATACATCTCTCCACCGTACCCAGGGACAG

FIGURE 18B

GAGTCAGGAAAAACCTGTCCTCCGTATAAGGAAGAGAAGGACGTGAAAACGCACCTAGTCCCAGAATGTCAG
AACAGTCTCACCTTCACTGGCACAAGCACATCAGCAACCCAGATGGCAGTCACCTTGCAGTAGATGACTTAAA
GAACGCTGAAGAGTCAAATAGGTCCAGATATTGGGATTCAAAGGAAGATGATGAAAGACAGACAGACTCTAAA
AAGAAGAAACTATTCCTCCAAGTTAAATAAAACAGATGTGATACATGGACAGGACAATCAGATGTTCAAAACACT
CAGCTAACAGTGGAAACGACAATATAGAAGGCACATTTCCTATCCCCTGGAAGAAACCAAATTACACGCTATT
CCCCGATGAAACGATCAATGTTGAAAACAATGAAGTCCAGAAGCTCGTCTATTCCCGGGAAAGAAAGCTGGTCG
GTGGGGTTAACCAAGGATGTAGAGTACAGTTCAATCACGGACCAGCAATTGACGACGGAATGGCAATGCCAAGTTCAA
AAGATCACGCGCTCTCATAGCACAGATATTCCCTACATTGTCGGAAGCTGCAGTGCAAGCTGAGCAAAAGAGCA
GTTGCAAGATATGCAAGATGAACACCATGTCGCTGAAGCAATTCTCGAATCCCTCGCTTGCCCTAACCTTACTG
ACAGAAATGGGATGGAAAACCTACTGTCGCTGAAGCCAGATCAAACCTTGGGATTCCCCTCTCAGGTCAAAAGT
TTACATGGACATCCTAGGAATGTGAAATCCATTAGGGAAAGTTAGACAGATCTGGACATGCCAGTAGTGTAAGCAG
CTTAGTAATTGTCGCTGGAATGACAGCAGAAGAAAAAAAGGTTAAGAAAGAGAAAGCTCCACAGAAACTGAATGCT
AGTCTGTTTGTCTTTAATTTTTAAACAGTCAGAAACCCACTAATGGGTGTCATCTGGCCATCCTAAA
CACATMTCCAATTCTAAAAACATTCCCTAAAAAAATTGGAAATTAGCTGATTTACAAATTAAATGCACT
AAAAGTAGTATTTGTTAGXATATGTTAGTAGGCTTAGTTTCAGTTGCAGTAGTATCAAATGAAAGTGATGATA
CTGTAACGAAGATAAAATTGGCTAATCAGTATAAGATTATAACATCTTTATTACTGAGGGCCACCAAATAGCCT
AGGAAGTGCCCTCGAGCACTGAAGTCACCATTAGGTCACTCAAGAAGTAAGCAACTAGCTGGGCACAGTGGCTCATG
CCTGTAATCCTAGCACTTGGGAGGCCAGAAAGATAGCTGAGTCCAGGAGTTGAGACCAGCCTGGCAAC
ATAGTGATACCCATCTTTAAAAAAAAAAAAAA

FIGURE 19

CTGAATCTCGTTCTCTCCCAGGGACCCCTCATTTCATATCCAGAAAATGTGATGCCACAGGTATCAGCGT
 CTGGATGCCACTTCACGTTAGCCACAAGTGAUTCAGTGGAAAGATCCAGAGTCAACAGAGGCTCGTCAGGAAGAT
GTTACAGAAAAGGTAGACCAAAAGGAGGAAGCTGGGAAAAAGAGGTGTGCGGAGACCAGATCAAAGGACCGACA
 AAGAGGAGGAACCACCAGCTGCTGCATCCCAGGCCAGGGTGGCGTCCAGGTGGCAGAGCAGCTAGGAACGCAAGG
 CCTGAACCTGGGCCAGACACCCCTGCTCTCCGCCATGGTCAACGACCCCTCCAGTACCTGCCTACTGTGGGCCA
 GGAGGTGGCCAAGTCTTGGCAGGCCGTGCCCCAGGCTGCTGCTGAGTTGGGGTGTCTTCTGCACCACCTCC
 TTTTGCTCTGGGTGTCTGTCTTCCCTATGGCTCCTTACTATTCCATATGCCAGACTCAGCCACCTCAGCCCT
 GTGCATTCTACTACAGGACCGACTGTGATTCCACCACCTCACTCTGCTCCTCCCTGTTGCCAATGTCTCGCT
 GACTAAGGGTGGACGTGATGGGTGCTGATGTATGGACAGCCGTACCTAGAGCTTGAGCTGCCAGAGT
 CCCCTGTGAATCAAGATTGGCATGTTCTGGTACCATTCCTGCTACACCAGAGGTGGCGAATCATCTCCACT
 TCTTCGCGTTCGGTGATGCTGCATTACCGCTCAGACCTGCTCCAGATGCTGGACACACTGGCTTCTAGCCTCCT
 GCTATTGGCTTGCAGAGCAGAAGCAGCTGCTGGAGGTGGAACCTACGCAGACTATAGAGAGAACTCGTACGTGC
 CGACCACGGAGCGATCATTGAGATCCACAGCAAGCGCATCCAGCTGTATGGAGCCTACCTCCGCATCCACGCGCAC
 TTCACTGGCTCAGATACTGCTATAACTCCGATGACCTGCGCCTCATAGGTGTTGCCAGCAACTCACCT
 CCTCAGCGTCATCGTCTTCAGCTACATGCAGTGGGTGCCCCGACACCGCTCTTTGC
 AGGTTAACATCCAAAAAGAGACAATTCCCGAAGGAAGTCAACGAAGGATCTGCTCATCAGCCAGGGCTGAA
 GGCCAGGAGGAGTCAACTCCGCAATCAGATGTTACAGAGGATGGTGAGAGCCCTGAAGATCCCTCAGGGACAGAGGG
 TCAGCTGTCCGAGGAGGAGAAACCAGATCAGCAGCCCTGAGCGGAGAAGAGGAGCTAGAGCCTGAGGCCAGTGATG
 GTTCAGGCTCCTGGAAAGATGCAGCTTGCTGACGGAGGCCAACCTGCCTGCTCCTGCTCCTGCTTCTGCC
 CCTGCTCTAGAGACTCTGGCAGCTGAAACCTGCTGGGGTGTCTCCGACAGCGCCCCACCTGCTCTAGTTCG
AAGAAAAGGGCAGACTCCTCACATTCCAGCACTTCCCACCTGACTCCTCTCCCTGTTTTCTCAATAAACT
ATTTGTGTCAGCTCAAAAAAAAAAAAAAA

FIGURE 20A

FIGURE 20B

GGGGACATTCCTCTGGAGAGAAGGAGATCAGGGCAACTCGGAGAGGGCTGCGAGCATTCCCTCCGGAGAG
GAAATCAGGGCGACCTGCACGCAGCTGCGTAGAGCCTGGAAGGGAAGTGGAAACCGACCGGACCGCCCTGCCCCCT
TCCCGGGATCACTTAATGAACCACGTGTTGACATCATGTTAACCTAACGACGTACAGATGATTCCGGATTGACA
AAATAACATTTGAGTATCCGATTGCCATCACCCCTACCCCCGAAATAGGACAACTCACTTCATTGACCAGGATGAT
CACATGGAAGGCAGCGCAGAGGCAGCTGTGGCTGCAGATTCTGTGTGGGTTCAAGCTATAAAACGCACCTC
CATCCGCCCTTCCCACAGCATTCTCCATCTAGATAGATGGTACTCTCAAAGGCCCTACCAGAGGAAACACGGC
CTACTGAGCGGACAGAATGATGCCAAATATTGCTTATGCTCTACATGGTATTGTAATGAATATCTGCTTTAATAT
AGCTATCATTCTTTCCAAAATTACTCTCTTATCTGGAATTAAATGAAATGAATTATCTGAATATAGG
AAGCATATGCCTACTTGTAAATTCTAACTTATGTTGAAGAGAAACCTCCGGTGTGAGATATAACAAATATATT
AATTGTGTCATATTAACCTCCCGGAATTC

FIGURE 21

GCATCTGGTTGTCAGATCCGAGAGGGCTCTGAAACTGCGGAGCGGCCACCGGACGCCCTCTGGAGCAGGTAGCAGCA
TGCAGCCGCCTCCAAGTCTGTGCGGACGCCCTGGTTGCGCTGGTCTTGCCTGCGCCTGTCGGATCTGGGA
 GAGGAGAGAGGCTCCCGCCTGACAGGGCCACTCCGCTTTGCAAACCGCAGAGATAATGACGCCACCCACTAAGAC
 CTTATGCCCAAGGGTCCAACGCCAGTCTGGCGCGTCGTTGGCACCTGCGGAGGTGCCTAAAGGAGACAGGACGG
 CAGGATCTCGCCACGCACCCTCCCCCTCCCCGTGCCAAGGACCCATCGAGATCAAGGAGACTTCAAATACATC
 AACACGGTTGTCTGCCTGTGTTGCTGGGATCATGGAACTCCACACTCTGAGAATTATCTACAAGAA
 CAAGTGCATGCAAACGGTCCAATATCTTGTGCTGGGAGACTGGCATTGGAGCTGAGATGTGTAAGCTGGTGCCTTCATA
 CAGAAAGCCTCCGTGGAAATCACTGTGCTGAGTCTATGTGCTTGAGTATTGACAGATATCGAGCTGTTGCTTCTG
 GAGTAGAATTAAAGGAATTGGGTTCCAAAATGGACAGCAGTAGAAATTGTTGATTGGTGGTCTCTGTGGTTC
 TGGCTGTCCTGAAGCCATAGGTTTGTATATAATTACGATGGACTACAAAGGAAGTTATCGGAATCTGCTTGCTT
 CATCCCGTTCAGAAGACAGCTTCATGAGTTTACAAGACAGCAAAGATTGGTGGCTGTCAGTTCTATTCTG
 CTTGCCATTGCCATCACTGCATTGTTATACACTAATGACCTGTGAAATGTTGAGAAAGAAAAGTGGCATGCA
 TTGCTTTAAATGATCACCTAAAGCAGAGACGGGAAGTGGCAAAACCGTCTTGCCTGGCCTGTCTTGCCCTC
 TGCTGGCTCCCTCACCTCAGCAGGATTCTGAAGCTCACTCTTATAATCAGAATGATCCAATAGATGTGA
 ACTTTGAGCTTCTGTTGGTATTGGACTATATTGGTATCAACATGGCTTCACTGAATTGCTGCATTAA
 ACCAATTGCTC
 TGTATTGGTGGAGAAAAGATTCAAAAATGCTTAAGTCATGCTTATGCTGCTGGTGC
 CAGTCCTGGAGGAAAAGCAGTCGTGCTAAAGTC
 AAAGCTAATGATCACGGATATGACA
 ACTTCCGTTCCAGTAA
 TAAATACAGCTCATCTTGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MKENYCLQAALVCLGMLCHSHAFAPERRGHLRPSFHGHHEKGKEGQVLQRSKRGWVNQFFVIEEYTGDPVLVGR
 HSDIDSGDGNIKYIISGEAGTIFVIDDKSGNIHATKLDREERAQYTLMAQAVDRDTNRPLEPPSEFIVKVQDIND
 NPPEFLHETYHANVPERSNVGTSVIQVTASDADDPTYGNSAKLVYSILEGQPYFSVEAQTCIIRTLPNMDREAKEE
 YHVVIQAKDMGGHMGGLSGTTKVTITLTDVNDNPPKFPQRLYQMSVSEAAVPGEEVGRVKAKDPDIGENGLVTYNIV
 DGDGMESFEITTDYETQEGVIKLKKPVDFETERAYSLKVEAANVHIDPKFISNGPFKDTVTVKISVEDADEPPMFLA
 PSYIHEVQENAAAGTVVGRVHAKDPDAANSPIRYSIDRHTDLDRFFTINPEDGFIKTTKPLDREETAWLNTVFAAE
 IHNRHQEAQPVVAIRVLDVNDNAPKFAAPYEGFICESDQTPLSNQPIVTISADDKDDTANGPRFIFSLPPEIIHNP
 NFTVRDNRDNTAGVYARRGGFSRQKQDLYLLPIVISDGGIIPPMSSNTLTIKVCGCDVNGALLSCNAEAYILNAGLS
 TGALIAILACIVILLVIVVLFTLRRQKKEPLIVFEEEDVRENIITYDDEGGGEEDTEAFDIATLQNPDGINGFIPR
 KDIKPEYQYMPRPGLRPAPNSVDVDDFINTRIQUEADNDPTAPPYDSIQIYGYEGRGSVAGSLSSLESATTDSLDDYD
 YLQNWGPRFKKLADLYGSKDTFDDDS

Signal sequence.

amino acids 1-22

Transmembrane domain.

amino acids 617-637

N-glycosylation sites.

amino acids 455-458, 540-543

Glycosaminoglycan attachment sites.

amino acids 83-86, 93-96

N-myristoylation sites.

amino acids 108-113, 215-220, 242-247, 246-251, 247-252, 399-404, 594-599, 599-604, 614-619, 618-623, 749-754, 753-758, 787-792

Cadherins extracellular repeated domain signatures.

amino acids 147-157, 256-266, 476-486

Cadherin cytoplasmic region.

amino acids 641-789

Cadherin domains.

amino acids 59-150, 164-259, 273-375, 388-479, 492-593

FIGURE 23

MYGKSSTRAVLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTWTWNFRPLDGGPEQFVF
YYHIDPFQPMMSGRFKDRVSDGNPERYDASILLWKLQFDDNGTYTCQVKNPPDVGVIGEIRLSVVHTVRFSEIHFL
ALAIGSACALMIIIVVVLFQHYRKKRWAERAHKVVEIKSKEEERLNQEKKVSVYLEDTD

Signal sequence.

amino acids 1-21

N-glycosylation sites.

amino acids 39-42, 118-121

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 205-208

N-myristylation sites.

amino acids 15-20, 119-124, 159-164

Immunoglobulin domain.

amino acids 40-125

FIGURE 24

MAEQQQGRELEAECPVCWNPFNNTFHTPKMLDCCHSFCVECLAHLSLVTPARRRLLCPLCRQPTVLASGQPVTDLPTD
TAMLTLRLLEPHHVILEGHQLCLKDQPKSRYFLRQPRVYTLGLGPQPGGQTGPPPDTASATVSTPILIPSHHSLREC
FRNPQFRIFAYLMAVILSVTLLLIFSIFWTKQFLWGVG

Transmembrane domain.

amino acids 162-182

N-glycosylation site.

amino acids 21-24

N-myristoylation site.

amino acids 68-73

Zinc finger, C3HC4 type, signature.

amino acids 32-41

FIGURE 25

MATAAGATYFQRGSLFWFTVITLSFGYYTWVVFWPQSIPYQNLGPLGPFTQYLVDHHHTLLCNGYWLA
YIALCKHKGITSGRAQLLWFLQTFFFGIASLTILIAKRKRQKQT

Transmembrane domain.

amino acids 57-77, 92-112

N-myristoylation site.

amino acids 87-92

Leucine zipper pattern.

amino acids 46-67

FIGURE 26

MASKIGSRRWMLQLIMQLGSVLLTRCPFWGCFSQLMLYAERAEARRKPDIPVPYLYFDMGAAVLCASFMSFGVKRRWFALGAALQLAISTYAAYIGGYVHYGDWLKVRMYSRTVAIIGGFLVLASGAGELYRRKPRSRSLQSTGQVFLGIYLICVAYSLQHSKEDRLAYLNHLPGGELMIQLFFVLYGILALAFLSGYYVTLAAQILAVLLPPVMLLIDGNVAYWHNTRVVEFWNQMKLLGESVGIFGTAVILATDG

Transmembrane domains.

amino acids 6-26, 51-71, 106-126, 138-158, 174-194, 201-221, 237-256

Glycosaminoglycan attachment site.

amino acids 125-128

N-myristoylation site.

amino acids 245-250

FIGURE 27

GRGSPLALLIRMKTLFGVWALLALILCPGVPEELFEVSIWPSQALVEFGQSLVCNCSTTCPDPGPQSGIETFLKKTQ
VDKGPKWKEFILLEDTENSILQCFFSCAGIQKDTSLGITVYQPPEQVILELQPAWVAVDEAFTVKCHVPSVAPLES
TLALLQGNQELHRKNFTSLAVASQRAEVIISVRAQKENDRCNSSCHAELDLSLQGGRLFQGSSPIRIVRIFEFSQSP
HIWVSSLLEAGMAETVSCEVARVFPKAKEVMFHMFLDQELSSLSWEGDTAWANATIRTMEAGDQELSCFASLGAME
QKTRKLVHSYSFPPPILELKESYPLAGTDINVTCGHLTSPSPTLRLQGAPDLPAGEPAWLLLTAREEDDGXNFSC
EASLVVQGQRLMKTTVIQLHILKPQLEESSCPGKQTWLEGMEHTLACVPKGNPAPALVCTWNGVVFDLEVPQKAT

Signal sequence.

amino acids 1-30

N-glycosylation sites.

amino acids 56-59, 169-172, 196-199, 285-288, 339-342, 382-385

N-myristoylation sites.

amino acids 3-8, 242-247, 335-340, 380-385, 425-430

Intercellular adhesion molecule (ICAM) homology.

amino acids 17-123

FIGURE 28

MLPRLLLICAPLCEPAELFLIASPSHPTEGSPVTLCKMPFLQSSDAQFQFCFFRDTRALGPGWSSSPKLQIAAMW
KEDTGSYWCEAQTMASKVLRSSRQINVHIPVSRPILMLRAPRAQAAVEDVLELHCEALRGSPPILYWFYHEDITLG
SRSAPSGGGASFNLSLTEEHSGNYSCEANGLGAQRSEAVTLNFTVPTGARSNHLTSGVIEGLLSTLGPATVALLFC
YGLKRKIGRRSARDPLRSLPQEFTRYLNSPTPGQLQPIYENVNVVSGDEVYSLAYYNQPEQESVAAETLGTHMED
KVSLDIYSRLRKANITDVEDAM

Signal sequence.

amino acids 1-16

Transmembrane domain.

amino acids 211-231

N-glycosylation sites.

amino acids 167-170, 177-180, 197-200, 322-325

Glycosaminoglycan attachment site.

amino acids 160-163

N-myristoylation sites.

amino acids 31-36, 82-87, 161-166, 163-168, 176-181, 187-192, 203-208, 212-217,
216-221

Amidation site.

amino acids 238-241

Immunoglobulin domains.

amino acids 31-88, 126-182

FIGURE 29

MLLWASLLAFAFAPVCGQSAAAHKPVISVHPPWTTFFKGERVLTLCNGFQFYATEKTTWYHRHYWGEKLTLPGNLEV
 RESGLYRCQARGSPRSNPVRLLFSSDSLILQAPYSVFEGDTLVLRCHRRKEKLTAVKYTWNGNILSISNKSWDLLI
 PQASSNNNGNYRCIGYGDENDVFRSNFKIIKIQELFPHPELKATDSQPTEGNSVNLSCTQLPPERSDTPLHFNFFR
 DGEVILSDWSTYPELQLPTVWRENCGSYWCGAETVRGNIHKHSPSLQIHVQRIPVSGVLLETQPSGGQAVEGEMLVL
 VCSVAEGTGDTTFWSHREDMQESLGRKTQRSLRRAELELPAIRQSHAGGYYCTADNSYGPVQSMVLNVTVRETPGNRD
 GLVAAGATGGLLSALLLAVALLFHCWRRRKSGVGFLGDETRLPPAPGPGESSHISICPAQVELQSLYVDVHPKKGDLV
 YSEIQTTQLGEEEANTSRTLLEDKDVSVVYSEVKTQHPDNSAGKISSKDEES

Signal sequence.

amino acids 1-15

Transmembrane domain.

amino acids 387-407

N-glycosylation sites.

amino acids 147-150, 209-212, 374-377, 478-481

Glycosaminoglycan attachment site.

amino acids 416-419

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 413-416

Tyrosine kinase phosphorylation site.

amino acids 457-463

N-myristoylation sites.

amino acids 15-20, 81-86, 89-94, 140-145, 163-168, 205-210, 257-262, 315-320,
 355-360, 382-387, 386-391, 391-396, 394-399, 395-400

Amidation site.

amino acids 332-335

Immunoglobulin domains.

amino acids 37-87, 116-169, 205-263, 303-361

FIGURE 30

MTVIRFFPAASATKRVLPVLRVSSPRTWNPNVPESPRI PAPRLPKRMSGAPTAAGAALMLCAATAVLLSAQGGPVQS
KSPRFASWDEMNVLAHGLLQLGQGLREHAERTSQLSALEERRLSACGSACQGTEGSTDPLAPESRVDPEVLHSLQT
QLKAQNSRIQQQLFHKVAQQQRHLEKQHLRIQHLQSQFGLLDHKHDHEVAKPARRKRLPEMAQPVDPAHNVSRLHRL
PRDCQELFQVGERQSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEK
VHSITGDRNSRLAVQLRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFSTWDQDHDLRRD
KNCAKSLSGGWWFGTCHSNLNGQYFRSI PQQRQLKKGIFWKTRGRYYPLQATTMLIQPMAAEAAAS

Transmembrane domain.

amino acids 49-69

N-glycosylation site.

amino acids 224-227

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 46-49, 118-121

N-myristoylation sites.

amino acids 50-55, 129-134, 341-346, 357-362

Fibrinogen beta and gamma chains C-terminal domain signature.

amino acids 396-408

Fibrinogen beta and gamma chains.

amino acids 231-447

FIGURE 31

MLLWVILLVLAPVSGQFARTPRPIIIFLQPPWTTVFQGERVLTCKGFRFYSPQKTKWYHRYLGKEILRETPDNILEV
QESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

Signal sequence.
amino acids 1-15

N-myristoylation site.
amino acids 89-94

FIGURE 32

MAGPRPSPWARLLAALISVLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDEMFRDRRCNTQAELLAAGCQRESI
 VVMESSFQITEETQIDTTLRSQMSPOGLRVRLPGEERHELEVFEPLESPVDDYILMDFNSNSMSDDLDNLKKMKGQ
 NLARVLSQLTSDYТИFGKFVDKSVPPQTDMPKPEKLKEPWNSDPPFSFKNVI
 SITEDVDEFRNKLQGERISGNLDA
 PEGGFDAILQTA
 VCTRDIGWRPDSTHLLV
 FSTESAFHYEADGANVLAGIMSRNDERCHLD
 TGTYTQYRTQDYPSPV
 TLV
 RLLAKHNI
 IPIFAVTNSY
 SYEKLHTYFPV
 SLLGV
 LQEDSSN
 IVELLEAFN
 RIRS
 NLD
 IALD
 S
 PRGLR
 TEV
 TSKMFQKTRTGSF
 HIRRGEV
 GIYQVQLRALEH
 VDGT
 HVCQL
 PEDQK
 GNIHLK
 PFS
 SDGLK
 MDAGI
 ICDV
 C
 CELQ
 KE
 VRSARCS
 FNGDF
 VCGQC
 VCSEG
 WSG
 QTC
 NC
 STG
 L
 SDI
 QPC
 CL
 REGED
 K
 PC
 S
 GR
 GEC
 QCG
 HC
 V
 CY
 GEG
 RY
 EG
 QF
 C
 EY
 D
 NFQC
 PR
 TSG
 F
 GL
 C
 N
 DR
 G
 R
 C
 M
 Q
 C
 V
 C
 E
 P
 G
 W
 T
 G
 P
 S
 C
 D
 C
 P
 L
 S
 N
 A
 T
 C
 I
 D
 S
 N
 G
 G
 I
 C
 N
 G
 R
 G
 H
 C
 E
 C
 G
 R
 C
 H
 C
 H
 Q
 Q
 S
 L
 Y
 T
 D
 T
 I
 C
 E
 I
 N
 Y
 S
 Y
 A
 I
 H
 P
 G
 L
 C
 E
 D
 L
 R
 S
 C
 V
 Q
 C
 Q
 A
 W
 G
 T
 G
 E
 K
 K
 G
 R
 T
 C
 E
 E
 C
 N
 F
 K
 V
 K
 M
 V
 D
 E
 L
 K
 R
 A
 E
 E
 V
 V
 V
 R
 C
 S
 F
 R
 D
 E
 D
 D
 C
 T
 Y
 S
 Y
 T
 M
 E
 G
 D
 G
 A
 P
 G
 P
 N
 S
 T
 V
 L
 V
 H
 K
 K
 D
 C
 P
 G
 S
 F
 W
 W
 L
 I
 P
 L
 L
 L
 L
 L
 P
 L
 L
 L
 L
 C
 W
 K
 Y
 C
 A
 C
 C
 K
 A
 C
 L
 A
 L
 L
 P
 C
 C
 N
 R
 G
 H
 M
 V
 G
 F
 K
 E
 D
 H
 Y
 M
 L
 R
 E
 N
 I
 M
 A
 S
 D
 H
 L
 D
 T
 P
 M
 L
 R
 S
 G
 N
 L
 K
 G
 R
 D
 V
 V
 R
 W
 K
 V
 T
 N
 N
 M
 Q
 R
 P
 G
 F
 A
 H
 A
 S
 I
 N
 P
 T
 E
 L
 V
 P
 Y
 G
 L
 S
 R
 L
 A
 R
 L
 C
 T
 E
 N
 L
 L
 K
 P
 D
 T
 R
 E
 C
 A
 Q
 R
 L
 Q
 E
 V
 E
 E
 N
 L
 N
 E
 V
 R
 Q
 I
 S
 G
 V
 H
 K
 L
 Q
 Q
 T
 K
 F
 R
 Q
 Q
 P
 N
 A
 G
 K
 Q
 D
 H
 T
 I
 V
 D
 T
 V
 L
 M
 A
 P
 R
 S
 A
 K
 P
 A
 L
 L
 K
 T
 E
 K
 Q
 V
 R
 A
 F
 H
 D
 L
 K
 V
 A
 P
 G
 Y
 Y
 T
 L
 T
 A
 D
 Q
 D
 A
 R
 G
 M
 V
 F
 Q
 E
 G
 V
 E
 L
 V
 D
 V
 R
 V
 P
 L
 F
 I
 R
 P
 D
 D
 E
 K
 Q
 L
 L
 V
 E
 A
 I
 D
 V
 P
 A
 G
 T
 A
 T
 L
 G
 R
 R
 L
 V
 N
 I
 T
 I
 I
 K
 E
 Q
 A
 R
 D
 V
 V
 S
 F
 E
 Q
 P
 E
 F
 S
 V
 S
 R
 G
 D
 Q
 V
 A
 R
 I
 P
 V
 I
 R
 R
 V
 L
 D
 G
 G
 K
 S
 Q
 V
 R
 T
 Q
 D
 G
 T
 A
 Q
 G
 N
 R
 D
 Y
 I
 P
 V
 E
 G
 E
 L
 F
 Q
 P
 G
 E
 A
 W
 K
 E
 L
 Q
 V
 K
 L
 L
 E
 L
 Q
 E
 V
 D
 S
 L
 L
 R
 G
 R
 V
 R
 F
 H
 V
 Q
 L
 S
 N
 P
 K
 F
 G
 A
 H
 L
 G
 Q
 P
 H
 S
 T
 I
 I
 I
 R
 D
 P
 D
 E
 L
 D
 R
 S
 F
 T
 S
 Q
 M
 L
 S
 Q
 P
 P
 P
 H
 G
 D
 L
 G
 A
 P
 Q
 N
 P
 N
 A
 A
 G
 S
 R
 K
 I
 H
 F
 N
 W
 L
 P
 S
 G
 K
 P
 M
 G
 Y
 R
 V
 K
 Y
 W
 I
 Q
 G
 D
 S
 E
 S
 E
 A
 H
 L
 D
 S
 K
 V
 P
 S
 V
 E
 L
 T
 N
 L
 Y
 P
 C
 D
 Y
 E
 M
 K
 V
 C
 A
 Y
 G
 A
 Q
 G
 E
 G
 P
 Y
 S
 S
 L
 V
 S
 C
 R
 T
 H
 Q
 E
 V
 P
 S
 E
 P
 G
 R
 L
 A
 F
 N
 V
 V
 S
 S
 T
 V
 T
 Q
 L
 S
 W
 A
 E
 P
 A
 E
 T
 N
 G
 E
 I
 T
 A
 Y
 E
 V
 C
 Y
 G
 L
 V
 N
 D
 D
 N
 R
 P
 I
 G
 P
 M
 K
 V
 L
 V
 D
 N
 P
 K
 N
 R
 M
 L
 L
 I
 E
 N
 L
 R
 E
 S
 Q
 P
 Y
 R
 Y
 T
 V
 K
 A
 R
 N
 G
 A
 G
 W
 G
 P
 E
 R
 E
 A
 I
 I
 N
 L
 A
 T
 Q
 P
 K
 R
 P
 M
 S
 I
 I
 I
 P
 I
 D
 P
 I
 P
 V
 D
 A
 Q
 S
 G
 E
 D
 Y
 D
 S
 F
 L
 M
 Y
 S
 D
 D
 V
 L
 R
 S
 P
 G
 S
 Q
 R
 P
 S
 V
 S
 D
 D
 T
 E
 H
 L
 V
 N
 G
 R
 M
 D
 F
 A
 F
 P
 G
 S
 T
 N
 S
 L
 H
 R
 M
 T
 T
 S
 A
 A
 Y
 G
 T
 H
 L
 S
 P
 H
 V
 P
 H
 R
 V
 L
 S
 T
 S
 S
 T
 L
 T
 R
 D
 Y
 N
 S
 L
 T
 R
 S
 E
 H
 H
 S
 T
 T
 L
 P
 R
 D
 Y
 S
 T
 L
 T
 S
 V
 S
 S
 H
 D
 S
 R
 L
 T
 A
 G
 V
 P
 D
 T
 P
 T
 R
 L
 V
 F
 S
 A
 L
 G
 P
 T
 S
 L
 R
 V
 S
 W
 Q
 E
 P
 R
 C
 E
 R
 P
 L
 Q
 G
 S
 V
 E
 Y
 Q
 L
 L
 N
 G
 E
 L
 H
 R
 L
 N
 I
 P
 N
 P
 A
 Q
 T
 S
 V
 V
 E
 D
 L
 L
 P
 N
 H
 S
 Y
 V
 F
 R
 V
 R
 A
 Q
 S
 Q
 E
 G
 W
 G
 R
 E
 R
 G
 V
 I
 T
 I
 E
 S
 Q
 V
 H
 P
 Q
 S
 P
 L
 C
 P
 L
 P
 G
 S
 A
 F
 T
 L
 S
 T
 P
 A
 P
 G
 L
 V
 F
 T
 A
 L
 S
 P
 D
 S
 L
 Q
 S
 W
 E
 R
 P
 R
 P
 N
 G
 D
 I
 V
 G
 Y
 L
 V
 T
 C
 E
 M
 A
 Q
 G
 G
 P
 A
 T
 A
 F
 R
 V
 D
 G
 D
 S
 P
 E
 S
 R
 L
 T
 V
 P
 G
 L
 S
 E
 N
 V
 P
 Y
 K
 F
 V
 Q
 A
 R
 T
 T
 E
 G
 F
 G
 P
 E
 R
 E
 I
 I
 T
 I
 E
 S
 Q
 D
 G
 G
 P
 F
 P
 Q
 L
 G
 S
 R
 A
 G
 L
 F
 Q
 H
 P
 L
 Q
 S
 E
 Y
 S
 S
 I
 T
 T
 H
 T
 S
 A
 T
 E
 P
 F
 L
 V
 D
 G
 P
 T
 L
 G
 A
 Q
 H
 L
 E
 A
 G
 G
 S
 L
 T
 R
 H
 V
 T
 Q
 E
 F
 V
 S
 R
 T
 L
 T
 S
 G
 T
 L
 S
 T
 H
 M
 D
 Q
 Q
 F
 Q
 T

Signal sequence.

amino acids 1-23

Transmembrane domain.

amino acids 711-731

N-glycosylation sites.

amino acids 327-330, 491-494, 579-582, 617-620, 695-698, 980-983, 1523-1526

Glycosaminoglycan attachment site.

amino acids 513-516

Tyrosine kinase phosphorylation sites.

amino acids 43-50, 401-408, 674-682, 1596-1604

N-myristoylation sites.

amino acids 24-29, 153-158, 234-239, 420-425, 477-482, 487-492, 560-565, 587-592, 588-593, 782-787, 1031-1036, 1089-1094, 1093-1098, 1383-1388, 1563-1568, 1613-1618, 1614-1619, 1635-1640, 1676-1681, 1719-1724, 1739-1744

Amidation sites.

amino acids 875-878, 974-977

Cell attachment sequence.

amino acids 1003-1006

EGF-like domain cysteine pattern signatures.

amino acids 479-490, 562-573

Integrins beta chain cysteine-rich domain signature.

amino acids 512-525, 590-603

Integrin, beta chain homology.

amino acids 37-455

Calx-beta domain.

amino acids 979-1084

Fibronectin type III domain.

amino acids 1127-1208, 1220-1310, 1458-1542, 1571-1658

FIGURE 33

MTLDRPGEATMLKFTVLLFCIRMSLGMTSIVMDPQPELWIESNYPQAPWENITLWCRSPSRISSKFLLLKDQTQM
 TWIRPSHKTQVFLIGALTESNAGLYRCCYWKETGWSKPSKVLEAPGQLPKPIFWIQAETPALPGCNVNILCHG
 WLQDLVFMFLKEGYAEPVDYQVPTGTMAIFSIDNLTPEDEGVYICRTHIQMLPTLWSEPSNPLKLVVAGLYPKPTLT
 AHPGPIMAPGESLNLRQCQGPIYGMTFALMRVEDLEKSFYHKKTIKNEANFFFQSLKIQDTGHYLCFYDASYRGSSL
 SDVLKIWVTDTPKTLARPSAVVQMGQNVSLRCRGPDGVGLALYKKGEDKPLQFLDATSIDNTSFFLNNVTYS
 DTGIYSCHYLLTWKTSIRMPHNTVELMVVDKPKPSLSAWPSTVFKLGKAITLQCRVSHPVLEFSLEWEERETFQR
 FSVNGDFIISNDGKGTGYSCSYRVEHPNMWSHRSEPLKLMGPAGYLWTWNYVLNEAIRLSLIMQLVALLLWVLWI
 RWKCRRLRIREAWLLGTAQGVVMLFIVTALLCCGLCNGVLIEETEIVMPTPKPELWAETNFPLAPWKNLTWCRSPS
 GSTKEFVLLKDGWIA TRPASEQVRAAFPLGALTQSHTGSYHCHSWEEMAVSEPESEALELVGTDILPKPVISASPT
 IRGQELQLRCKGWLAGMGFALYKEGEQE PVQQLGAVGREAFTT IQRMEDKDEGNYS CRTHTEKLPFKWSEPEPLEL
 VIKEMYPKPFKWTASPVVTPGARVTFCNCSTPHQHMSFILYKDGSEIASSDRSWASPGASAHHFLIISVGIGDGGNY
 SCRYYDFSIWSEPSDPVELVVTEFYPKPTLLAQPGPVVFPGKSVILRCQGTFOGMRFALLQEGAHVPLQFRSVSGNS
 ADFLLHTVGAEDSGNYSCIYYETTMSNRGSYLSMPLMIWVTDTPKWLFAEPSSVPMQGVTLWCRGPVHGVGYI
 LHKEGEATSMQLWGSTSNDGAFPITNISGTSMGRYSCCYHPDWTSSIKIQPSNTLELLVTGLLPKPSLLAQPGPMVA
 PGENMTLQCQGELPDSTFVLLKEGAQEPLEQQRPSGYRADFWMPAVRGEDGDIYSCVYILDSTPFAASNHSDSLEIW
 VTDKPPKPSLSAWPSTMFKLGKDTLQCRGPLPGVEFVLEHDGEEAPQQFSEDGDFVINNVEGKGIGNYSCSYRLQA
 YPDIWSEPSDPLELVGAAGPVAQECTVGNIVRSSLIVVVVVALGVVLAIEWKKWPRLTRGSETDGRDQTIALEECN
 QEGEPGTPANSPSSTSQRISVELPVPI

Signal sequence.

amino acids 1-28

Transmembrane domains.

amino acids 517-537, 555-575, 1261-1281

N-glycosylation sites.

amino acids 53-56, 338-341, 374-377, 381-384, 607-610, 747-750, 798-801, 846-849, 939-942, 986-989, 1027-1030, 1082-1085, 1147-1150, 1223-1226

Tyrosine kinase phosphorylation sites.

amino acids 287-295, 1125-1132

N-myristoylation sites.

amino acids 102-107, 145-150, 195-200, 254-259, 305-310, 336-341, 349-354, 388-393, 480-485, 555-560, 573-578, 656-661, 709-714, 746-751, 792-797, 814-819, 828-833, 840-845, 844-849, 845-850, 897-902, 901-906, 933-938, 938-943, 953-958, 984-989, 1015-1020, 1030-1035, 1130-1135, 1222-1227, 1276-1281, 1315-1320

Immunoglobulin domains.

amino acids 51-108, 145-201, 241-298, 336-394, 434-486, 605-662, 696-752, 792-851, 888-944, 984-1040, 1080-1136, 1176-1228

FIGURE 34

MAPEPAPGRTMVPALVMLGLVAGAHGDSKPVFIKVPEDQTGLSGGVASFVCQATGEPKPRITWMKKGKKVSSQR
 FEVIEFDDGAGSVLRIQPLRVQRDEAIYECTATNSLGEINTSAKLSVLEEEQLPPGFPSIDMGPQLKVVEKARTATM
 LCAAGGNPDPEISWFKDFLPVPATSNRIKQLRSGALQIESSEESDQGKYECVATNSAGTRYSAPANLYVRVRRVA
 PRFSIPSSQEVMPPGSVNLTCVAVGAPMPYVKWMMGAEELTKEDEMPVGRNVLELSNVVRSANYTCVAISLGMIE
 ATAQVTVKALPKPPIDLVVTETTATSVTLTWDGNSEPVYYGIQYRAAGTEGPFQEVGDGVATTRYSIGGLSPFSEY
 AFRVLAVNSIGRGPSEAVRARTGEQAPSSPPRRVQARMLSASTMLVQWEPPENGLVRGYVYYTPDSRRPPNAW
 HKHNTDAGLLTVGSLLPGITYSLRVLAFTAGDGPSPTIQVKTQQGVPAQPADFQAEVESDTRIQLSWLLPPQER
 IIMYELVYWAAEDEDQQHKVTFDPTSSYTLEDLKPDTLYRFQLAARSMDMGVGVFTPTIEARTAQSTPSAPPQKVMCV
 SMGTTVRVSWVPPPADSNGVITQYSVAHEAVDGEDRGRHVVDGISRREHSSWDLVGLEKWTLEYRVWVRAHTDVGPG
 PEPPSPVLRVTDDEVPSGPPRKVEVEPLNSTAVHVYWKLPVPSKQHQQIRGYQVTVRLENGEPRGLPIQDVMLAEA
 QWRPEESEDYETTISGLTPETTYSVTVAAYTTKGDGARSKPKIVTTGAVPGRPTMMIYSTAMNTALLQWHPPKELP
 GELLYGRLQYCRADEARPNTIDFGKDDQHFTVTGLHKGTTYIIFRLAAKNRAGLGEFEKEIRTPEDLPSGFQONLHV
 TGLTTSTTELAWDPVLAERNGRIISYTVVFRDINSQOELQNITTDTRFTLTGLKPDTTYDIKVRRAWTSKGSGPLSP
 SIQSRTMPVEQVFAKNFRVAAAMKTSVLLSWEVPSYKSAVPKIYLNGQSVEVDGHSMRKLIAQLQPNTEYSFVLM
 NRGSSAGGLQHLVSI RTAPDLLPHKPLPASAYIEDGRFDLSMPHVQDPSLVRWFYIVVVPIDRVGGSMLTPRWSTPE
 ELELDELLEAIEQGGEERQRRRRQAERLKPYVAAQLDVLPEFTLGDKNYRGFYNRPLSPDLSYQCFVLASLKEPM
 DQKRYASSPYSDEIVVQVTPAQQQEPEMLWVTGPVLAVILILIVIAILLFKRKRTHSPSKDEQSIGLKDSLH
 SSDPVEMRRLNYQTPGMRDHPPIPITDLADNIELRKANDGLKFSQEYESIDPGQQFTWENSLEVNKPKNRYANVIA
 YDHRSVILTSIDGVPGSDYINANYIDGYRKQNAYIATQGPLPETMGDFWRMVWEQRTATVVMTRLEEKSRVKCDQY
 WPARGTETCGLIQVTLLDTVELATYTVRTFALHKGSSEKRELROFQFMAWPDHGVPEYPTPILAFLRRVKACNPLD
 AGPMVVHCSAGVGRGCFIVIDAMLERMKHEKTVDIYGHVTCMRSQRNYMVQTEDQYVFIEALLEATCGHTEVPA
 RNLYAHIQKLGQVPPGESVTAMELEFKLLASSKAHTSRFISANLPCNKFKNRLVNIMPYELTRVCLQPIRGVEGSDY
 INASFLDGYRQQKAYIATQGPLAESTEDFWRMLWEHNSTIIVMLTKLREMGREKCHQYWPAERSARYQYFVVDPMAE
 YNMPQYILREFKVTDARDGSRTIRQFQFTDWPEQGVPKTGEGFIDFIGQVHKTKEQFGQDGPITVHCSAGVGRGTV
 FITLSIVLERMRYEGVVDMFQTVKTLRTQRPAMVQTEDQYQLCYRAALEYLGSDFDHYAT

Signal sequence.

amino acids 1-29

Transmembrane domain.

amino acids 1262-1282

N-glycosylation sites.

amino acids 117-120, 250-253, 295-298, 721-724, 966-969, 1696-1699, 1731-1734
 cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 71-74, 1286-1289

Tyrosine kinase phosphorylation sites.

amino acids 97-105, 773-780, 1229-1237, 1687-1694

N-myristoylation sites.

amino acids 22-27, 44-49, 47-52, 214-219, 246-251, 305-310, 368-373, 442-447,
 470-475, 481-486, 739-744, 818-823, 881-886, 926-931, 1081-1084, 1301-1306,
 1349-1354, 1362-1367, 1399-1404, 1468-1473, 1551-1556, 1688-1693, 1790-1795,
 1807-1812, 1842-1847, 1847-1852

Amidation site.

amino acids 69-72

Leucine zipper pattern.

amino acids 1262-1283

Myb DNA-binding domain repeat signature 1.

amino acids 1151-1159

Tyrosine specific protein phosphatases active sites.

amino acids 1546-1558, 1837-1849

Immunoglobulin domains.

amino acids 47-109, 149-209, 246-300

Fibronectin type III domains.

amino acids 319-401, 413-500, 512-594, 606-696, 708-809, 821-904, 915-1000

Protein-tyrosine phosphatase homology.

amino acids 1375-1606, 1664-1897

FIGURE 35

MRRAALWLWL CALALSLQLALPQIVATNLPPEDQDGSGDDSDNFSGSGAGALQDITLSQQT PSTWKDTQLLTAIPTS
PEPTGLEATAASTSTLPAGEGPKEGEAVVLPEVEPGLTAREQEATPRPRETTQLPTTHQASTTATTAQEPATSHPH
RDMQPQGHETSTPAGPSQADLHTPHTEDGGPSATERAAEDGASSQLPAAEGSGEQDFTFETSGENTAVVAVEPDRRN
QSPV DQGATGASQGLLDRKEVLGGVIAGGLVGLIFAVCLVGFM LYRMKKDEGSYSLEEPKQANGGAYQKPTKQEEFYA

Signal sequence.

amino acids 1-22

Transmembrane domain.

amino acids 252-272

N-glycosylation site.

amino acids 43-46

Glycosaminoglycan attachment sites.

amino acids 45-48, 47-50

Tyrosine kinase phosphorylation site.

amino acids 279-286

N-myristoylation sites.

amino acids 46-51, 82-87, 183-188, 238-243, 241-246, 254-259, 255-260, 259-264,
263-268

Syndecans signature.

amino acids 276-289

Syndecan domain.

amino acids 3-308

FIGURE 36

MKIFLPVLLAALLGVERASSLMCFSCLNQKSNLYCLKPTICSDQDNYCVTVSASAGIGNLVTFGHSLSKTCSPACPI
PEGVNNGVASMGISSCQSFLCNFSAADGGLRASVTLLGAGLLSLLPALLRFGP

Signal sequence.

amino acids 1-20

Transmembrane domain.

amino acids 108-128

N-glycosylation site.

amino acids 99-102

N-myristoylation sites.

amino acids 14-19, 58-63, 80-85, 89-94, 105-110, 106-111, 117-122

u-PAR/Ly-6 domain.

amino acids 21-100

FIGURE 37

MDGKKCSVWMFLPLVFTLFTSAGLWIVYFIAVEDDKILPLNSAERKPGVKHAPYISIAGDDPPASCVFSQVMNMAAF
LALVVAVLRFIQLKPKVLPWLNISGLVALCLASFGMTLLGNFQLTNDEEIHNVGTSLTFGFTLTCWIQAALTAKV
NIKNEGRRVGIPRVIILSASITLCVVLYFILMAQSIHMYAARVQWGLVMCFLSYFGTFAVEFRHYRYEIVCSEYQENF
LSFSESLSEASEYQTDQV

Transmembrane domains.

amino acids 7-27, 65-85, 97-117, 131-151, 166-186

N-glycosylation site.

amino acids 100-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-7

N-myristoylation sites.

amino acids 132-137, 140-145, 199-201

Amidation sites.

amino acids 2-5, 159-162

FIGURE 38

MELALLCGLVVMAGVIPIQGGILNLNKMVKQVTGKMPILSYWPYGC~~H~~GLGGRGQPKDATDWCCQTHDCCYDHLKTQ
GCGIYKDYYRYNFSQGNIHCSDKG~~W~~CEQQLCACDKEVAFCLKRNLDTYQKRLRFYWRPHCRGQTPGC

Signal sequence.
amino acids 1-17

N-glycosylation site.
amino acids 89-92

N-myristoylation sites.
amino acids 20-25, 45-50, 93-98, 140-145

Phospholipase A2 histidine active site.
amino acids 63-70

Phospholipase A2 aspartic acid active site.
amino acids 108-118

Phospholipase A2 homology.
amino acids 21-145

FIGURE 39

MYIRVSYDTKPDSSLHLMVKDWQLELPKLLISVHGLQNFEMQPKLKQVFGKGLIKAAMTTGAWIFTGGVSTGVISH
 VGDALKDHSSKSRRVCAIGIAPWGIVENKEVLGVKDVTRYQTMSNPLSKLSVLNNNSHTHFILADNGTLGKYGAEV
 KLRLLEKHISLQKINTRLGQGVPLVGLVVEGGPNVVSIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGGI
 INESLREQLLVTIQKTFNYNKAQSHQLFAIIMECMKKELVTVFRMGSEGQQDIEMAILTALLKGTNVSAFDQLSLA
 LAWNRVDIARSQIFVFGPHWTPLGLSAPPTDSKATEKEKKPPMATTGGGRGKGKGGKGKVKVEEVEETDPRKIELL
 NWVNALEQAMLDALVLDRVDFVKLLIENGVNMQHFLTIPRLLELYNTRLGPPNTHLLVRDVKSNLPPDYHISLID
 IGLVLEYLMGGAYRCNYTRKNFRFLYNNLFGPKRPKALKLGMEDDEPPAKGKKKKKKKEEEIDIDVDDPAVSRFQ
 YPFHELMVWAVLMKRQKMAFLWQRGEESMAKALVACKLYKAMAHESESSDLVDDISQDLDNNSKDFGQLALELLDQ
 SYKHDEQAMKLLTAKLAVAALKHDFIAHTCSQMLLTDMWMGRRLMRKNPGLKIVMIGILLPPTILF
 LEFRTYDDFSYQTSKENEDGKEKEEENTDANADAGSRKGDEEENEHKKQRSIPIGTKICEFYNAPIVKFWFTISYLG
 YLLLNFNYVILVRMDGWPSLQEWVISYIVSLALEKIREILMSEPGKLSQKIKVWLQEWNIITDLVAISTFMIGAILR
 LQNQPYMGYGRVIYCVDIIFWYIRVLDIFGVNKYLGPYVMMIGKMMIDMLYFVVIMLVVLMSFGVARQAILHPEEK
 SWKLARNIFYMPYWMIFYGEVFADQIDLYAMEINPPCGENLYDEEGKRLPPCIPGAWLTPALMACYLLVANILLVNLL
 IAVFNNFFEVKSISNQVWKFQRYQLIMTFHDPVLPMMIILSHIYIIIMRLSGRCRKREGDQEERDRLKLF
 DEELKRLHEFEEQCVQEHFREKEDEQQSSSDERIRVTSERVENMSMRLEEINERETFMKTSQTVDLRLAQLEELSN
 RMVNALENLAGIDRSDLIQARSRASSECEATYLLRQSSINSADGSLYRYHFNGEELLFEDTSLSTSPGTGVRKKT
 SFRIKEEKDVKTHLVECPQNSLHLSLGTSTSATPDGSHLAVDDLKNAEESKLGPDIGISKEDDERQTDISKKEETISP
 SLNKTDTIHGQDKSDVQNTQLTVETTNIEGTISYPLEETKITRYFPDETINACKTMKRSRSFVYSRGRKLVGGVNQDV
 EYSSITDQQLTTEWQCQVQKIRSHSTDIPYIVSEAAVQAEQKEFADMQDEHHVAEAIPIPRLSLTITDRNGMEN
 LLSVKPDQTLGFPSSLRSKSLHGHPRNVKSIQGKLDRSGHASSVSSLIVSGMTAEKKVKKEKASTETEC

Transmembrane domains.

amino acids 759-779, 828-848, 857-877, 893-913, 976-996, 992-1012, 1031-1051

N-glycosylation sites.

amino acids 133-136, 144-147, 233-236, 298-301, 478-481, 601-604, 635-638, 638-641, 830-833, 1006-1009, 1121-1124, 1312-1315

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 1228-1231

Tyrosine kinase phosphorylation sites.

amino acids 697-704, 891-898

N-myristoylation sites.

amino acids 35-40, 53-58, 68-73, 69-74, 102-107, 211-216, 229-234, 296-301, 473-478, 728-733, 747-752, 1166-1171, 1259-1264, 1268-1273, 1319-1324

Amidation sites.

amino acids 362-365, 513-516, 968-971, 1374-1377

Ion transport protein homology.

amino acids 789-1005

FIGURE 40

MSTEKVDQKEEAGEKEVCGDQIKGPDKEEEPPAASHGQGWRPGGRAARNARPEPGARHPALPAMVNNDPPVPALLWA
QEVGQVLAGRARRLLLQFGVLFCTILLWWVSFVLYGSFYYSYMPTVSHLSPVHFYYRTDCDSSTTSICSPVANVS
LTKGGGRDRVLMYQGPYRVTLELELPESPVNQDLGMFLVTISCYTRGGRIISTSSRSVMLHYRSDLQMLDTLVFSSL
LLFGFAEQKQLLEVELYADYRENSYVPTTGAIIEIHSKRIQLYGAYLRIAHFTGLRYLLYNFPMTCAFIGVASNFT
FLSVIVLFSYMQWWGGIWPRHRFSLQVNIRKRDNSRKEVQRRISAHQPGPEGQEESTPQSDVTEDGESPEDPSGTE
GQLSEEKPDQQPLSGEEELEPEASDGGSWEDAALLTEANLPAPAPASASAPVLETLGSSEPAGGALRQRPTCSSS

Transmembrane domains.

amino acids 94-114, 134-154, 292-312, 308-328

N-glycosylation sites.

amino acids 152-155, 306-309

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 350-353

Tyrosine kinase phosphorylation site.

amino acids 240-248

N-myristoylation sites.

amino acids 44-49, 81-86, 96-101, 302-307, 361-366

Leucine zipper patterns.

amino acids 84-105, 91-112

FIGURE 41

MGPPPLPLLLLLLPPRVLPAAPSSVPRGRQLPGRLGCLLEEGLCASEACVNDGVFGRCQKVPAMDFYRYEVSPV
 ALQRLRVALQKLSGTGFTQDDYTQYVMDQELADLPKTYLRRPEASSPARPSKHSVGSERYSREGGAALANALRRH
 LPFLEALSQAPASDVLARTHTAQDRPPAEGDDRFSESIILTYVAHTSALTYPPGPRQLREDLLPRTLGQLQPDELSP
 KVDGVDRHHLMAALSAAYAQRPPAPPGEGLSLEPQYLLRAPSMPRPLLAPAAPQKWPSPLGDSEDPSSTGDGARIH
 TLLKDLQRQPAEVRLSGLELDGMAELMAGLMQGVDHGVARGSPGRAALGESGEQADGPKATLRGDSFPDDGVQDDD
 DRLYQEVRHLSATLGGLLQDHGSRLPGALP FARPLDMERKKSEHPESSLSEEETAGVENVKSQTYSKDLLGQQPH
 SEPGAAAFGELQNQMPGPSKEEQLSPAGAQEALSDGLQLEVQPSEEEARGYIVTDRDPLRPEEGRRLVEDVARLLQV
 PSSAFADVEVLGPATFKVSANQNVTTEDVEKATVDNKDKLEETSGLKILQTGVGSKSCLKFLPPQAEQEDSTKFI
 ALTLVSLACILGVLLASGLIYCLRHSSQHRLKEKLSGLGGDPGADATAAAYQELCRQRMATRPPDRPEGPHTSRISSV
 SSQFSDGPIPSARSSASSWSEEPVQSNMDISTGHMILSYMEDHLKNKNRLEKEWEALCAYQAEPNSSFVAQREEN
 VPKNRSLAVLTYDHSRVLLKAENSHSHSDYINASPIMDHDPRNPAYIATQGPLPATVADFWQMVWESGCVVIVMLTP
 LAENGVRQCYHYWPDEGSNLYHIYEVNLVSEHIWCEDFLVRSFYLKNLQTNETRTVTQFHFLSWYDRGVPSRSLL
 DFRRKVNKCYRGRSCPIIVHCSGDAGRSPTYVLIDMVLNKMAKGAKEIDIAATLEHLRDQPGMVQTKEQFEFALTA
 VAEEVNAILKALPQ

Signal sequence.

amino acids 1-21

Transmembrane domain.

amino acids 616-636

N-glycosylation sites.

amino acids 564-567, 760-763, 774-777, 898-901

Glycosaminoglycan attachment sites.

amino acids 90-93, 652-655

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 137-140, 425-428

N-myristoylation sites.

amino acids 44-49, 47-52, 91-96, 144-149, 338-343, 342-347, 346-351, 490-495, 628-633, 634-639, 655-660, 659-664, 852-857, 915-920, 948-953, 987-992

Amidation site.

amino acids 525-528

Cell attachment sequence.

amino acids 372-374

Tyrosine specific protein phosphatases active site.

amino acids 943-955

Protein-tyrosine phosphatase.

amino acids 770-1004

FIGURE 42

MQPPPSLCRALVALVLACGLSRIWGEERGFPPDRATPLLQTAEIMTPPTKTLWPKGNSASLARSLAPAEVPKGDRTAGSPPRTISPPPCQGPIEIKETFKYINTVVSCLVFVLGIIGNSTLLRIIYKNKCMRNGPNILIASLALGDLHIVIDIPINVYKLLAEDWPFGAEMCKLVPFIQKASVGITVSLCALSIDRYRAVASWSRIKGIGVPKWTAVEIVLIWVVSVV LAVPEAIGFDIITMDYKGSYLRICLLHPVQKTAFMQFYKTAKDWWLFSFYFCLPLAITAFFYTLMTCEMLRKYSGMQIALNDHLKQRREVAKTVFCLVLVFALCWPLHLSRILKLTLYNQNDPNRCELLSFLLVLDYIGINMASLNSCINPIA LYLVSKRFKNCKSCLCCWCQSFEEKQSLEEKQSLKFKANDHGYDNFRSSNKYSSS

Transmembrane domains.

amino acids 1-19, 101-121, 137-157, 177-197, 216-236, 275-295, 323-343, 362-382

N-glycosylation sites.

amino acids 59-62, 119-122

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 302-305

Tyrosine kinase phosphorylation site.

amino acids 424-430

N-myristoylation sites.

amino acids 57-62, 115-120, 170-175, 306-311, 371-376

7 transmembrane receptor homology.

amino acids 118-386